D I S S E R T A T I O N

zur Erlangung des akademischen Grades

Doctor rerum naturalium (Dr. rer. nat.) im Fach Psychologie

Risks, rewards, and rationality: How knowing that you probably won't hit the jackpot affects your judgments and decisions

eingereicht an der Lebenswissenschaftlichen Fakultät der Humboldt–Universität zu Berlin

von **Christina Leuker, M.Sc.**

Präsidentin der Humboldt-Universität zu Berlin
Prof. Dr. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin
Prof. Dr. Bernhard Grimm

Gutachter/innen:

1. Prof. Dr. Neil Stewart
2. Prof. Dr. Bettina von Helversen
3. Prof. Dr. Ralph Hertwig

Tag der mündlichen Prüfung: 13.11.2018

Berlin, im Juli 2018

Summary

It is a truism that nearly everyone prefers larger rewards to smaller rewards. However, as nearly everyone also knows, large rewards rarely occur as a windfall. Instead, risks and rewards are typically inversely related. At the same time, theories of decision making typically treat risks and rewards as two *independent* attributes that determine the value of an option and ultimately choice (Savage, 1954; von Neumann & Morgenstern, 1944). Thus, there is a disconnect between the types of choice environments people experience outside the lab and the types of choice environments theories have been derived from. It is possible that the same theories and conclusions—that have been derived from decision making studies with independent, uncorrelated risks and rewards—apply to cases in which there is a systematic relationship between risks and rewards. However, theories of adaptive cognition make different predictions; namely, that the environment is represented or reflected in the mind, and these representations of the environment in the mind systematically affect how it operates (e.g. Anderson, 1991; Brunswik, 1944; Gibson, 1979). In this dissertation, I theoretically and empirically examined how the mind adapts to risk–reward structures and how the link between risks and rewards systematically affects judgments and decisions.

The first chapter provides a broad theoretical overview of the key concepts this dissertation covers. Most of the experiments subsequently presented in this dissertation pit choice behavior in environments with more representative risk–reward structures—in which an inverse relationship between risks and rewards is present—against choice behavior in less representative environments in which there is no relationship between risks and rewards. What varies across experiments and chapters is the type and degree of uncertainty decision makers face. In Chapter 2, I show how people exploit risk–reward structures in decisions under uncertainty; that is, decisions in which probabilities are unavailable to the decision maker or difficult to ascertain. In these situations, consistent with the principle of *vicarious functioning* of different cues (Brunswik, 1943), people can infer the probabilities of events directly from the payoffs when risks and rewards are found to be correlated. Chapter 3 shows that risk–reward structures are a context variable that shapes how people evaluate options in decisions under risk. Surprising options that did not fit surrounding risk–reward structures were linked to longer response times and an increase in pupil size, particularly when options were "surprisingly good"—i.e., when they offered a high payoff and a high probability in an environment with otherwise inversely related risks and rewards. Chapter 4 addresses how risk–reward structures affect decisions under risk in general. A computational modeling account showed that risk–reward structures do not change (subjective) preferences in choices, even if they are among very high windfall amounts. Instead, risk–reward structures affect how people accumulate evidence in risky choice. Specifically, inversely related risks and rewards promote satisficing whereas uncorrelated risks and rewards promote maximizing. That is, an uncorrelated environment seems to "breed" more careful, maximizing decision makers. In Chapter 5, I provide an example of how risk–reward structures affect decision making in the wild, and in a case in which losses are possible. Specifically, I show that some individuals use very high pay as a cue to infer the potential risks a clinical trial poses; and that these risk assessments in turn influence how ethically inappropriate people find clinical trials.

Taken together, people exploit their intuition that there is usually "no free lunch" across a wide range of decisions. Evidence comes from a combination of behavioral experiments, eye-tracking data, computational modeling, and a replication and extension of a large online survey in which the link between risks and rewards had been overlooked. The work complements earlier research showing that risks and rewards are tied in the environment and because of this link are also tied in the human mind. In sum, this work suggests that risk–reward priors should not be blindly assumed away, and challenges assumptions on who is considered rational and why.

Zusammenfassung

"Wer nicht wagt, der nicht gewinnt!" Laut dieser Redensart müssen wir Risiken eingehen, um die großen Gewinne oder Ziele im Leben zu erreichen—sei es im Beruf, in Beziehungen, oder in finanziellen Fragen. Mit anderen Worten, in vielen Entscheidungsszenarien sind Gewinnhöhe und Gewinnwahrscheinlichkeit antikorreliert: Je höher der Gewinn—desto geringer die Wahrscheinlichkeit (Pleskac & Hertwig, 2014). Im Gegensatz dazu wird in bedeutenden Entscheidungstheorien angenommen, dass Gewinnhöhe und Gewinnwahrscheinlichkeiten *unabhängig* voneinander variieren. Ausgehend von Theorien der "Adaptiven Kognition" kann dies zum Problem werden, denn es lässt sich vorhersagen, dass der Mensch statistische Strukturen aus seiner Umwelt mental repräsentiert und diese Strukturen kognitive Prozesse systematisch beeinflussen (u.A. Anderson, 1991; Brunswik, 1944; Gibson, 1979).

Wie lernen Menschen die Kovariation zwischen Gewinnen und Wahrscheinlichkeiten aus der Umwelt? Wie beeinflussen solche Strukturen menschliche Entscheidungen und Urteile? Diese Fragen stehen im Fokus dieser Dissertation. Im 1. Kapitel gebe ich einen breit gefächerten Überblick über die Konzepte und Theorien, auf denen diese Arbeit aufbaut. Anschliessend präsentiere ich verschiedene Experimente, in denen ich Entscheidungen in repräsentativen Kontexten—in denen Gewinne und Wahrscheinlichkeiten negativ miteinander korrelierren—mit Entscheidungen in weniger repräsentativen Kontexten—in denen Gewinne und Wahrscheinlichkeiten unkorreliert oder positiv korrliert sind—vergleiche. Über Experimente und Kapitel hinweg variiert der Grad und die Art der Unsicherheit, denen die Entscheider (Versuchspersonen) ausgesetzt sind. Im 2. Kapitel zeige ich experimentell, dass eine Korrelation zwischen Gewinnhöhe und Gewinnwahrscheinlichkeit flexibel erlernt werden kann. In anschließenden Entscheidungen unter Unsicherheit (keine Gewinnwahrschenlichkeit gegeben) schätzten die Versuchspersonen die Gewinnwahrscheinlichkeit gemäß der zuvor erlernten Korrelation anhand der Gewinnhöhe ein. Im 3. Kapitel zeige ich experimentell, dass Versuchspersonen durch die Korrelation zwischen Gewinnhöhe und Gewinnwahrscheinlichkeit Erwartungen aufbauten: War eine Option "überraschend", z.B. da sie einen hohen Gewinn mit einer hohen Wahrscheinlichkeit verspricht, wurde eine solch überraschend gute Optionen länger evaluiert. Dies hat einen direkten Bezug zu nichtexperimentellen Kontexten, in denen sich solche Optionen oft als "zu gut um wahr zu sein" herausstellen. Im 4. Kapitel beschäftige ich mich mit der Frage der externen Validität—also der Generalisierbarkeit von (bisherigen) Untersuchungsergebnissen. In Experimenten zu Entscheidungen unter Risiko sind Gewinne und deren Wahrscheinlichkeiten oft unkorreliert und bilden daher nur einen Bruchteil der Entscheidungskontexte, die Menschen außerhalb des Labors vorfinden, ab. Mithilfe eines kognitiven Modells zeige ich, dass eine Struktur, in der hohe Gewinne sich als unwahrscheinlich erweisen, zu einer schnelleren, einfacheren Entscheidungsstrategie führen, als eine unkorrelierte Struktur. Im 5. Kapitel zeige ich, wie das von vielen Menschen angenommene "No free Lunch" Prinzip sich auf die Beurteilung von Risiken im Bereich klinischer Studien auswirkt: Hohe Vergütungen für klinische Studien können dazu führen, dass die Studie als riskanter und schlussendlich unethischer wahrgenommen wird.

Zusammengefasst lässt sich sagen, dass Menschen sich in unterschiedlichsten Entscheidungssituationen von der Intuition, dass man sprichwörtlich nichts im Leben geschenkt bekommt, leiten lassen—einer Intuition, mit der sie oft richtig liegen. Empirische Evidenz hierfür bieten die in dieser Dissertation präsentierten Verhaltensstudien in Kombination mit Blickbewegungsaufnahmen, Computational Modeling und einer Replikationsstudie in denen der (subjektive) Bezug zwischen Bezahlung und Wahrscheinlichkeit von Risiken zuvor nicht in Betracht gezogen wurde. Die Arbeit ergänzt vorherige Forschung, indem sie zeigt, dass Gewinne und Wahrscheinlichkeiten nicht nur in der Welt systematisch miteinander verbunden sind, sondern auch in der menschlichen Kognition. Dies wirft ein neues Licht auf eine alte Frage: Wer entscheidet wann "rational", und warum?

Contents

1 | Introduction

"If you hold a lungful of air you can survive in the total vacuum of space for about thirty seconds. However, what with space being the mindboggling size it is, the chances of getting picked up by another ship within those thirty seconds are two to the power of two hundred and seventy-six thousand seven hundred and nine to one against."

Douglas Adams, The Hitchhiker's Guide to the Galaxy

There is a ludicrously small chance of surviving if you get thrown off a space ship. So exploring space seems pretty dangerous—but, as the saying goes, *no risk, no reward.*[1] Whether a decision maker is weighing the costs and benefits of space travel or considering whether to consume food past its expiration date: risks and rewards, or payoffs and probabilities, are the key ingredients that determine the subjective value of an option and ultimately choice. Risks and rewards guide the decisions of human and nonhuman animals as they forage for habitats, food, relationships—and of course also decisions among many other, less important things such as money. Often, risks and rewards are treated as two independent building blocks that determine what rational action to take (von Neumann and Morgenstern, 1944; Savage, 1954; Sutton and Barto, 1998). However, psychologists have acknowledged what decision makers "in the wild" and ecologists have known for a long time: Risks and rewards are not independent from one another at all (Pleskac and Hertwig, 2014). In many domains, the larger rewards that human and nonhuman animals desire are unlikely to occur, and there are a few good reasons for why this is the case.

One of the reasons for this trade-off between reward magnitude and likelihood rests on how people and animals distribute themselves in relation to resources. As they forage, they will distribute themselves proportional to the total amount of resources in each patch, forming what behavioral ecologists call an *ideal free distribution* of competitors (IFD, Fretwell and Lucas, 1970). For instance, when bees forage for pollen, larger rewards—areas with a higher concentration of flowers—attract more bees per unit time (Ohashi and Yahara, 2002; Dreisig, 1995). This has consequences for bees subsequently arriving at the scene—for whom the large rewards become more unlikely. As they observe an initial group of bees already sampling from the largest rewards, the later-arriving bees (at least ideally) distribute themselves among the other patches offering smaller rewards. Because there will be fewer competitors at these patches, there will be a greater chance of obtaining the smaller rewards than if the bees go to the patch with larger resources. That is,

[1] In the Hitchhiker's Guide to the Galaxy, traveling space is also the superior option: Traveling the galaxy might kill Arthur and Ford, but staying on Earth definitely would have killed them.

an ideal free distribution of competitors can produce a risk–reward structure (Pleskac, Conradt, Leuker, & Hertwig, 2018). The ideal free distribution is not limited to bees, but extends to many different animal species as they forage for resources (for an extensive review see Davies et al., 2012; Kennedy and Gray, 1993). Ideal free distributions also occur for humans. Examples include students foraging for rewards in behavioral studies (Goldstone and Ashpole, 2004; Goldstone et al., 2005; Sokolowski et al., 1999), or searching for jobs (Krueger, 1988; Holzer et al., 1991). In sum, the ideal free distribution of competitors emerges as a group-level phenomenon as animals and humans seek to maximize their individual resource consumption (Pleskac et al., prep), and as the IFD is present in many different environments, a risk-reward relationship is implied in many environments as well.[2]

Another primary reason that risks and rewards are tied for human decision makers is the way in which many modern human choice environments are designed. For instance, the risk-reward structure can often be traced back to the forces of the marketplace. People want to *buy* options that offer high rewards with high probabilities at low costs, and other people want to *sell* options that offer low rewards at high costs. To make a transaction attractive for both sellers and buyers, the set of options available in many markets are pushed toward a fair bet. Consequently, as the expected payoffs and costs of gambles are pushed toward their fair price, the probability of a payoff will be inversely related to the magnitude of the payoff. Examples of this range from the gambles people play at the casino, such as roulette in which they need to trade off reward magnitude for odds of winning; to the gambles scientists play when submitting scientific journals in which they need to trade off impact factors for acceptance rates (Pleskac and Hertwig, 2014).

Rewards are also unlikely in social domains, where people often *share* rewards even though this means a smaller payoff for themselves. Sometimes people share for their own interest—as to not forego payoffs for themselves. This is a robust result from the ultimatum game, in which a so–called proposer can choose how to split a set monetary amount. A responder can accept or decline the proposed offer. The caveat in the standard form of this game is that if the responder chooses to reject the offer, both players will get nothing (Güth et al., 1982). Typically, very small offers will be rejected and very large offers will be accepted (Hoffman et al., 1996; Güth, 1995). Again, from the perspective of the proposer, an inverse relationship between the magnitude of the reward and its probability emerges (Pleskac and Hertwig, 2014). In the ultimatum game, proposers (need to) anticipate the minimal share that a responder would reject—which implies the proposer looses their payoff as well. But even without the fear of loosing own shares, people tend to share with others and thereby reduce what they keep to themselves. This has been shown experimentally in the dictator game: A so–called dictator is given a windfall resource to allocate between himself or herself and another player (the recipient), who has no right to reject the offer. Typically, dictators will allocate around 30% of the payoff to the recipient, very few give nothing (Engel, 2011; Liu et al., 2016). In these more altruistic forms of giving up rewards, smaller rewards for oneself may trade off against the larger benefit gained from cooperation (Bowles, 2006; Boyd, 2006).

While risk–reward structures are pervasive in human and nonhuman choice situations, they have been largely ignored in the study of judgment and decision making. Risks and rewards, or probabilities and

[2]While the IFD is pervasive in many natural domains, it has some preconditions that need to be met: One precondition is that the environment is competitive—i.e. resources are limited (e.g. there is no IFD for air). Moreover, competitors need to be *ideal*—e.g. they need to be able to detect the most promising patches (Pleskac et al., prep). The more an environment deviates from these preconditions, the less likely it is to have a risk–reward structure.

payoffs, are the two fundamental components any rational decision maker needs to consider to make good choices. In fact, the canonical approach to studying decision making is based on the idea that all states of the world can be re–described as subjective risks and rewards (Meder et al., 2013). These risks and rewards can then be combined into an expected utility (Savage, 1954; von Neumann and Morgenstern, 1944; Luce and Raiffa, 1957). In this view, risks and rewards are treated as independent—with the constraint that people are usually probed about their choices among nondominated options (i.e., $p_A > p_B$ and $x_A < x_B$, where p refers to probabilities/risks and x refers to payoffs/rewards for options A and B).

The use of nondominated options does create a *local* risk–reward structure where risks and rewards are inversely related, but this relationship only extends to the (typically two) options under consideration. The risk–reward relationship focused on in this dissertation is the *global* risk–reward relationship that exists across all possible gambles, for instance across all possible lottery tickets a customer could purchase or all the possible bets a player could select from at the roulette table. For such choice environments, a good understanding if (and if yes, how) a decision maker uses the relationship between risks and rewards, is lacking. It is possible that risk–reward structures do not impact choice at all. In this case, the same theories and conclusions—that have been by and large derived from decision making studies with independent, uncorrelated risks and rewards—apply to cases in which there is a systematic relationship between risks and rewards. However, theories of adaptive cognition make different predictions; namely, that the environment is represented or reflected in the mind, and these representations of the environment in the mind systematically affect how it operates (e.g., Anderson and Schooler, 1991; Brunswik, 1944; Gibson, 1979; Stewart et al., 2006; Ungemach et al., 2011; Shepard, 1987).

In this dissertation, I theoretically and empirically examined how the mind adapts to risk–reward structures and how the link between risks and rewards systematically affects judgments and decisions. Before moving on to experimental data, I outline the theoretical basis for my work. Many people before me have studied how risks and rewards *are* or *should be* combined into the value of an option, and what this implies for choice and human rationality. I briefly review this work by drawing upon different conceptualizations of uncertainty. I then address the question of how risk–reward structures can theoretically impact decision making under risk and uncertainty and discuss the role of representative experimental designs as an additional motivation for the experiments reported in this dissertation. The introduction concludes with a brief overview of the chapters and their main findings.

Decision making under risk and uncertainty

Uncertainty has many dimensions, among which a frequently used distinction is the one between *aleatory* and *epistemic uncertainty* (Hacking, 1975). *Aleatory uncertainty* stems from the environment with its inherent randomness, or inherent stochasticity. This uncertainty is irreducible (Figure 1A): Even in a simple lottery such as roulette—in which the outcome is either a pre-defined gain or an outcome of 0—the outcomes are probabilistic and there is no amount of learning that can change that uncertainty.

Epistemic uncertainty stems from the mind, and is therefore somewhat under the decision maker's control. People can decide to sample more information if the stakes of a decision are high (Hau et al., 2008), or discontinue search (exploit) if they have found a high reward (such as in the exploration–exploitation tradeoff, see Sutton and Barto, 1998). People can (and maybe should) also sample more if the situation has not been explored at all and therefore lack own experience, and/or social information or information from similar alternatives. Yet, actors often have practical limitations such as limited computational power, time and knowledge that prevent them from infinitely sampling information about a given choice situation.

Fortunately, infinitely sampling information in each and every choice situation may not be necessary. Instead, people are often able to generalize from past experiences to new choice situations. One way to do this is would be via sets of previously learned strategies from which they can choose in a given situation to which these strategies are applicable (Todd et al., 2012; Gigerenzer et al., 1999; Rieskamp and Otto, 2006). In order for such a generalization to work, there should be a fit between the mind's mechanisms and the structure of the environment (Todd and Gigerenzer, 2007). Taking this perspective implies that a third form of uncertainty, *systemic* uncertainty, can arise from a mismatch between the mechanism or tool the mind selects and the actual structure of the environment. Systemic uncertainty complements the standard dualistic view of uncertainty according to which uncertainty can unambiguously be ascribed either to the actor or the environment (Kozyreva et al.). In this dissertation, the focus is on how decision makers would in principle be *able to* and empirically *do* reduce systemic uncertainty as a function of the risk–reward environment they find themselves in—that is, how they may exploit risk–reward structures when they are present and refrain from doing so when no structure is given or can be expected.

Reducing systemic uncertainty by applying the right strategy in the right situation is challenging if the environment changes. The past can only be a good predictor of the future if the environment is relatively stable. Figure 1B depicts the possible changes using a simplified urn model. A decision maker may have intuited the possible payoffs and probabilities in a given choice situation (ii). Over time, the probabilities associated with a known set of payoffs—i.e. the risk–reward structure in a given environment—can vary (i), the case I am concerned with in this dissertation. In other cases, both the possible payoffs and probabilities change (iii).

Different sources of uncertainty and how they interact with changes in the environment can lead an actor to face a variety of decision making scenarios (Figure 1C), in which the degree of uncertainty varies. I will briefly review what is known about decision making under different degrees of uncertainty and outline how risk–reward structures may impact them. Probably the most scientifically studied scenario are decisions under risk. In decisions under risk, the probabilities of the outcomes in a situation are known and measurable (they range from $0 - 1$). The values 0 and 1 refer to the situation of certainty, and they have sometimes been referred to as a "degenerate form of risk" (Luce and Raiffa (1957), p. 13). In decisions under risk and certainty, a decision maker exactly knows just *how* stochastic or random an environment is. An event is least predictable if all possible payoffs are equiprobable, and it becomes more predictable the more one of the possible payoffs approaches the two extremes, 0 or 1.

These properties of decisions under risk—known payoffs and known probabilities—are also present in monetary lotteries, in which people are asked to choose among two or more options with explicitly stated risks and rewards (e.g. "Do you prefer a €30 for sure or €40 with a probability of .8?").[3] Be

[3]The "monetary lotteries" approach is nearly as old as the study of decision–making itself. There are at least three reasons for the popularity of monetary gambles. First, they serve as an abstraction of real decisions. The idea is that all possible states of the world can be described in terms of (subjective) payoffs, or rewards, and probabilities. Second, monetary gambles provide a normative framework for rational choice. For instance, a decision maker should always choose the option

it in an experimental monetary lottery or in the "wild": Which option should a decision maker choose when facing two or more alternatives with known payoffs and known probabilities? In 1654, an exchange of letters on gambling problems between French mathematicians Blaise Pascal and Pierre Fermat gave rise to the concept of mathematical expectation (Hacking, 1975). A decision under risk was thought to be rational if it maximized the decision maker's expected value. If given a choice between two or more gambles with different expected values, the objectively better choice is the gamble with the greater expected value. Formally, a simple gambles' EV can be obtained by multiplying its payoff x by its probability p ($EV = p \times x$). If the gamble offers multiple payoffs (or losses) with a unique probability each, the expected value is the sum of these payoff–probability combinations.

However, people's choices often violate(d) the idea of pure mathematical expectation. One of the most famous examples of such a violation is the St. Petersburg Paradox that Daniel Bernoulli 1954 used to motivate the development of expected utility theory. In the same paper, Bernoulli also used another gamble to motivate his new theory. In this case, he asked the reader to imagine "a very poor fellow" gets a lottery ticket that offers a ".5 chance of winning 20,000 ducates, otherwise nothing" (p. 24). If he is offered 9,000 ducates for this ticket—a sure win—, he would probably go ahead and sell the ticket; even though the expected value of keeping the ticket would have been higher. Thus, a gambles' *objective* worth may deviate from a gambles' *subjective* worth. In Bernoulli's words: "The determination of the value of an item must not be based on the price, but rather on the utility it yields... There is no doubt that a gain of one thousand ducats is more significant to the pauper than to a rich man though both gain the same amount" (p. 24). This is acknowledged in expected utility theory, according to which a rational decision maker should pick the option offering the highest expected *utility* (Bernoulli, 1954). Formally, expected utility theory can be implemented as an adjustable parameter that determines how subjective and objective worth map onto each other ($EU = p \times x^{\alpha}$, where the α parameter determines the degree of diminishing marginal utility of the outcomes).

But even expected utility theory could not account for all choice patterns observed as people chose among monetary gambles that allowed them to—in principle—compute expected utilities (e.g. in the Allais paradox that I describe in detail later). Some of the violations have been summarized as the fourfold pattern of risk attitudes: decision makers have been found to be risk-seeking over low-probability gains and risk-averse over high-probability gains. Just like the "very poor fellow" in Bernoulli's example, they often tend to value certain outcomes over probabilistic ones when there is a lot at stake. This changes in the loss domain, where people tend to be risk-averse over low-probability losses and risk-seeking over high-probability losses. Arguably, because it can explain these preferences and many other violations of rational choice theory, cumulative prospect theory has become the most influential descriptive model in the expectation tradition (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). Formally, expected utility is extended in prospect theory by the addition of a probability weighting function that describes the

that offers the highest subjective value. His or her subjective value can be inferred from his or her choices among gambles. Third, empirical violations of normative theory observed in monetary lotteries serve as benchmarks for theories of choice. In other words, if a theory could not account for how people chose among gambles it had lost its generalizeability, often leading to modifications of it. The most prominent theories of decision making have been shaped and reshaped based on the how people choose among monetary gambles (Bernoulli, 1954; von Neumann and Morgenstern, 1944; Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). For these and other reasons, gambles are considered to play an as big of a role for decision scientists as the "fruitfly [plays for] genetics" (p. 137, Lopes, 1983)

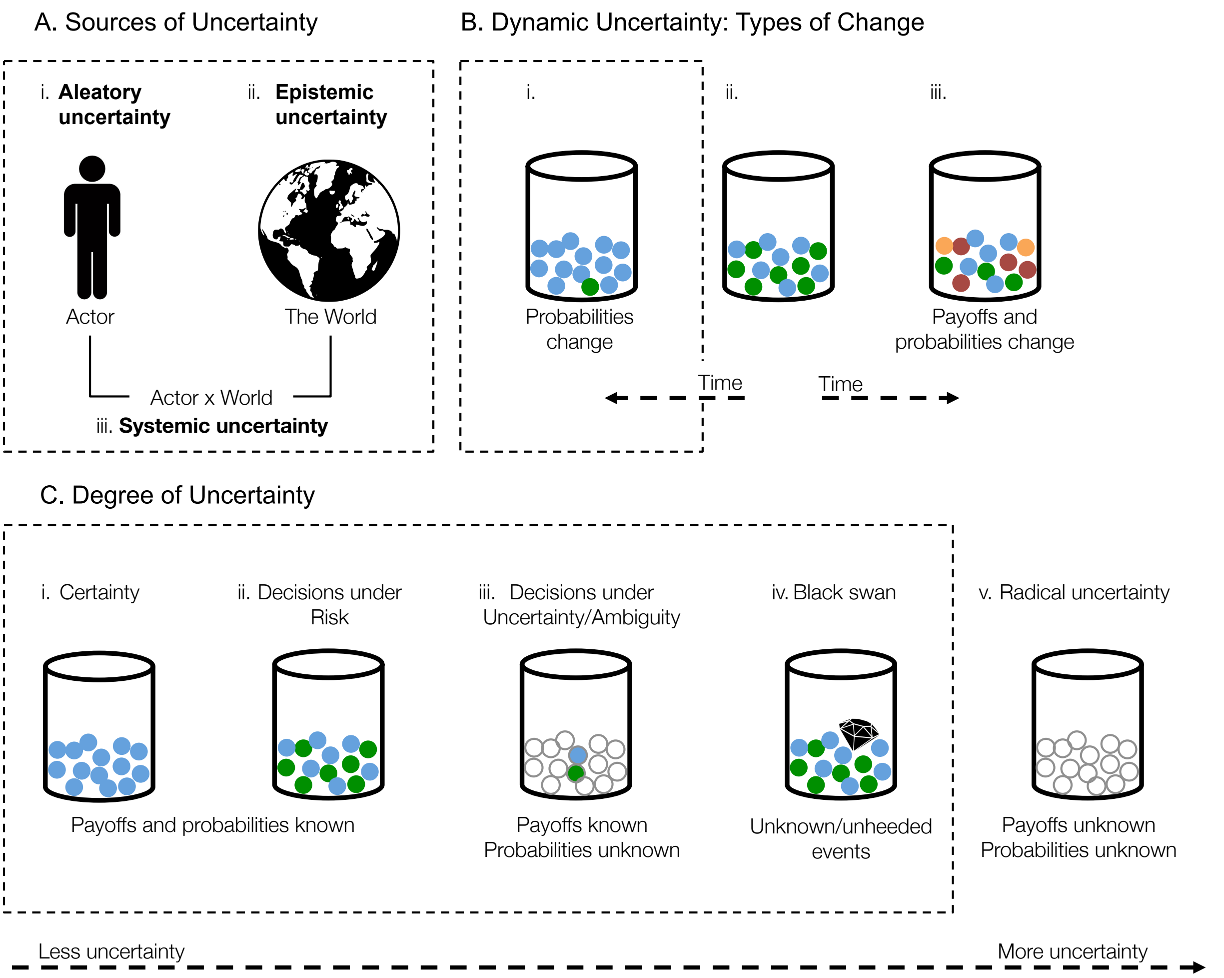


Figure 1. Decision making under various forms of uncertainty, illustrated by an urn model. **(A)** *Epistemic uncertainty* stems from the boundedly rational actor— for instance, from his or her lack of experience, time, or computational power. With greater epistemic uncertainty, a decision maker has less (subjective) knowledge but this can be remedied through learning or experience. *Aleatory uncertainty* stems from the world itself— some environments will always have more randomness than others. Aleatory uncertainty is greater when the probabilities are close to .5 in a simple two–outcome event and lower when the probabilities are to the extremes, 0 and 1.0. *Systemic uncertainty* can arise from the question of which mechanism or tool the mind should select in a given environment. Systemic uncertainty is the mutual product of environmental unpredictability and the actor's epistemic limitations (Kozyreva et al.). **(B)** The structure of the world may change in different ways. Some states of the world could become more likely than before (changing probabilities); or in some cases both the possible states of the world (payoffs) and their probabilities may change. If these changes occur, the probability estimates of the current situation (center) are of little use. **(C)** Degrees of uncertainty. In situations of certainty (i) and risk (ii), the probabilities are known. In decisions under uncertainty or ambiguity (iii), the payoffs are known but the probabilities are not. In a 'black swan' situation, a rare but highly impactful event is unknown to the decision maker or not represented in the decision situation (iv). The 'black swan' event can either be negative or positive. Under radical uncertainty, neither payoffs nor probabilities are known (v). Boxes indicate the elements of uncertainty addressed in this dissertation. Figure adapted from Meder et al. (2013).

extent to which objective probabilities are distorted when transformed into subjective decision weights.[4] From the concept of mathematical expectation to prospect theory, it seems that the goal was always to explain peoples choices in simple monetary gambles via modifications to existing theory. The most recent and most prominent one, prospect theory, has also been shown to account for various phenomena "in the wild", especially in the monetary domain (e.g. stock market behavior, horse betting, markets for consumer

[4]In the gain domain. If the set of choices one wants to account for also includes losses, there is at least one more parameter that captures the extent to which losses are amplified relative to gains (loss aversion).

goods Camerer, 2000). Linking prospect theory to nonlaboratory phenomena as a deductive approach has merits, yet all of these theories of risky choice have by and large been derived from environments in which risks and rewards are uncorrelated. Therefore, existing accounts of how people make risky decisions may be incomplete (see next two sections, and Chapter 4).

Can expected utility theory and its refined variants help in understanding decisions under uncertainty (iii), or are these accounts also incomplete? In many situations, probabilities are unknown, unknowable or at least difficult to ascertain (Knight, 1921; Luce and Raiffa, 1957; Pleskac et al., 2015). When probabilities are not available or unknown to the decision maker, expected utilities are also difficult to assess. One straightforward solution to circumvent missing probabilities is to simply replace them with subjective probability estimates (subjective expected utility theory, Savage, 1954). Under complete uncertainty, a decision maker could assume all outcomes to be equally likely (Laplace, 1776, quoted in Hacking, 1975, p. 132). If objective probabilities can simply be replaced by subjective ones, the distinction between decisions under risk and decisions under uncertainty becomes arbitrary. In essence, all uncertainties can be reduced to *risks*, at least for "rational man" (Savage, 1954, also see Ramsay as cited in Ellsberg, 1961). The subjective probabilities need to be consistent—but not necessarily plausible (Savage, 1954; Pleskac et al., 2015). For example, if there are two possible states of the world, A and B, and the subjective probability of A is .8 then the subjective probability of B needs to be .2—irrespective if state A is a decision maker's (very optimistic) subjective estimate of winning the lottery. Moreover, choices should match subjective probabilities. This solution, replacing objective probabilities with subjective ones, would have been quite a happy ending to the vexing problem of decision making under uncertainty (Figure 1C, iii). However, yet again, evidence from simple monetary gambles speaks against this view. Decisions under uncertainty can be represented as a monetary gamble in which the probabilities are partially or fully occluded (Tymula et al., 2012; van den Bos and Hertwig, 2017; Ellsberg, 1961), and people's choices in these gambles suggest that they are ambiguity averse. For instance, most people given a choice between two otherwise equivalent options—one in which the probability information is given and the other in which it is missing—avoid the option with missing probability information (Camerer and Weber, 1992). Suppose a decision maker is given a choice between these two lotteries:

$$
\begin{aligned}
A^* &: \text{Win \$100 if a red marble is drawn from urn I with 50 red marbles and 50 black marbles.}\\
B &: \text{Win \$100 if a red marble is drawn from urn II with 100 marbles with an unknown proportion}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{of red and black marbles.}
\end{aligned}
$$

People generally prefer A to B (Camerer and Weber, 1992; Ellsberg, 1961). This is still consistent with subjective expected utility theory—people may just have estimated the number of red marbles to be lower than 50 in lottery B (e.g. 40). Now consider a second set of options, referring to the same urns as above:

A* : Win \$100 if a black marble is drawn from urn I with 50 red marbles and 50 black marbles.

B : Win \$100 if a black marble is drawn from urn II with 100 marbles with an unknown proportion of red and black marbles.

In this scenario, people still prefer that a marble is drawn urn I—the urn with a known chance of winning. If subjective probability estimates would have explained choices in the first set of options, people should prefer urn II in the second case, as the color of the winning marble is re-assigned to black (their estimate for black should be $100 - 40 = 60$). This pattern of preferences suggests that uncertainty cannot simply be reduced to a probability, as subjective expected utility theory would assume (Savage, 1954). In other words, people do not simply behave "as if" they assign quantitative likelihoods to uncertain events, but they seem to dislike ambiguous probabilities due to the very fact that they *are* unknown, rather than a dislike of randomness or the stochasticity of a choice situation, per se (i.e. a 50% chance of winning).

Of course, there may be degrees of uncertainty. Even if probabilities are not explicitly stated, actors could invest time to learn probabilities from data and ultimately restrict the range of plausible values. If an environment is relatively stable, past experience can be a good predictor of the future (Figure 1B). Thus, people may sometimes be able to transition from a decision that looks like a decision under uncertainty to one that is more like a decision under risk (probabilities known). As such, one solution to the missing probability information is for people to sample information from the choice environment, either from the options themselves, from similar options, or social information, to form an impression about the likelihood of the different events (see for example Barron and Erev, 2003; Denrell, 2007; Hertwig and Erev, 2009; Pleskac, 2008; Weber et al., 2004).[5] Conversely, if an actor is completely uninformed about both the payoffs and the probabilities of an event, epistemic uncertainty is highest ("radical uncertainty"). But in many instances and across a range of environments, people may have already learned that high payoffs are unlikely. This means has consequences for decisions under uncertainty (e.g. the Ellsberg urns): After learning about risks and rewards being inversely related, people may infer probabilities of winning high payoffs (e.g. \$100) to be low, and consequently prefer an urn that offers \$100 with a probability of 50% (also see Chapter 2).

Another special case of uncertainty is the "black swan" scenario, in which a rare but highly impactful event is unknown to the decision maker and has not entered his or her mental representation of possible outcomes. Such events can easily be missed and many people will never experience them. Examples of positive black swans may be sudden high windfall profits (e.g. from winning the lottery or from investing in stocks that later highly exceed ones expectations). Such events are likely psychologically distinct because they are surprising given the risk–reward structures people are accustomed to (also see Chapter 3).

In sum, uncertainty comes in many forms. Conceptualizing choice situations in terms of these classes of uncertainty helps appreciating that "empirically studying decision making under uncertainty is anything

[5]But, to further complicate things, it has also been argued that uncertainty cannot be reduced to risk even with an infinite amount of data, as the data–generating process is ever–changing (Levitt and List, 2009)

but trivial" (Meder et al., 2013, p. 258). Generally, decisions under uncertainty may consist of an *attribution* to the type of uncertainty and different *degrees* of uncertainty each component entails (Fox and Ülkümen, 2011). For each situation, several questions arise: How much uncertainty stems from the decision maker (epistemic uncertainty), the environment (aleatory uncertainty), or an interaction between the two (systemic uncertainty)? To what extent can a decision maker reduce some of the uncertainties? In the case of risk–reward structures, a "candidate" type of uncertainty that both can be reduced and a decision maker may aim to reduce may be systemic uncertainty. The next section outlines some prerequisites and consequences of reducing systemic uncertainty in risk–reward environments.

Vicarious functioning and the mind–environment fit

Systemic uncertainty can be reduced in different ways. Some solutions to systemic uncertainty can be situated in the mind (Kozyreva et al.). For example, it is sometimes assumed that a decision–maker has a repertoire of well–defined strategies that he chooses among by considering the expected costs and benefits of each strategy (Rieskamp and Otto, 2006; Payne et al., 1988). The selection process could be a conscious process of applying a meta-strategy or an unconscious selection triggered by experience. In both cases, the mind has to be sufficiently sensitive to how the environment is structured, and have an impression of how well a particular strategy will work. Other solutions to systemic uncertainty can be situated in the demands of the environment—for example, a decision maker can only sample more information if he or she has enough time and cognitive resources to do so. Yet other solutions to systemic uncertainty can be situated in the structure of the environment—for example, a decision maker can ignore some of the cues in a choice problem if these cues are interrelated (Brunswik, 1952). From this perspective, environments with systematic relationships between their variables can be exploited by the decision maker, particularly when he or she lacks explicit information, time, or experience. In such kind environments in which reliable statistical regularities exist, exploiting their relationships (by, for instance, inferring one attribute from the other) still leads to good choices (Gigerenzer et al., 1999).

The reason for this is the fit between how the mind uses cues from the environment and the relationships between cues in the environment itself. One way to understand the mind–environment fit is by turning to what Brunswik called probabilistic functionalism (Brunswik, 1952, 1955; Pleskac and Hertwig, 2014; Dhami et al., 2004). Functionalism refers to the idea that people act in order to achieve a particular goal—for instance, finding a city they would like to live in or maximizing their payoffs in a risky choice experiment. The attribute value of the goal—the distal criterion—is often not directly accessible to them but there are many proximal cues that are probabilistically related to the attribute value. When seeking for a city to live in, one may turn to proximal cues that are more readily available, such as knowledge of the size of the city, the job prospects and living costs. The more predictive a criterion is of the underlying attribute value, the higher its ecological validity (see left hand side of Figure 2).

In decisions under risk, the distal criterion a decision maker seeks to maximize is the expected value or expected utility. The proximal cues to achieve this are the payoffs and probabilities an option offers. As Pleskac and Hertwig (2014) have shown, payoffs and probabilities, or risks and rewards, are inversely related across many domains in the environment. In the lens model framework, the risk–reward relationship

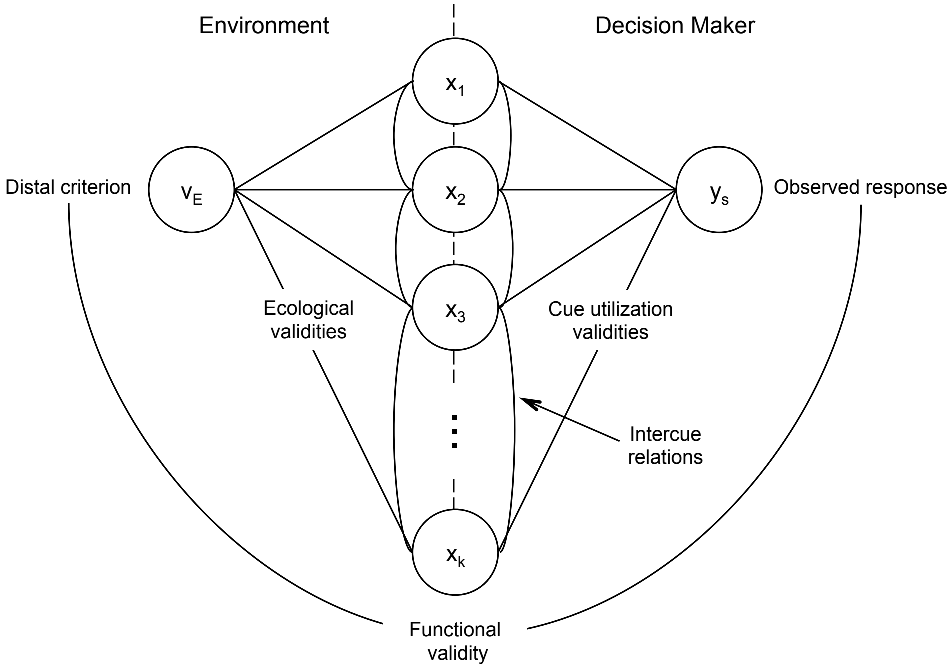


Figure 2. Brunswik's lens model (adapted from Brunswik, 1952 and Pleskac & Hertwig, 2014).

can be conceptualized as an *intercue relation* between the key elements that determine the value of an option, and ultimately choice (see center in Figure 2). This redundancy between risk and rewards offers an interesting solution to decisions under uncertainty, where payoffs may be known but probabilities are not (Knight, 1921; Luce and Raiffa, 1957). One of the proximal cues, the probability, is missing.

One way to attain probabilities in decision making under uncertainty is to exploit the (typically inverse) relationship between risks and rewards in form of a heuristic—the risk-reward heuristic—and infer the probability directly from the payoff (Pleskac and Hertwig, 2014). Brunswik (1955) referred to this as *mutual substitutability* or *vicarious functioning* of cues. The better the inferred probability matches the structure of the environment, the better it can be subsequently combined into the options' value (by weighing the payoff by it). Ultimately, such a match between the mind's inference and the environmental structure can result in a high degree of functional validity (connecting arc between distal criterion and observed response, see Figure 2). Indeed, Pleskac and Hertwig (2014) found evidence that the inverse relationship between risks and rewards seems to be represented in, and exploited by, peoples' minds. In a survey, they offered participants a gamble offering a chance to win $\$x$ at the cost of \$2, and asked them to estimate the probability of winning \$2.5, \$4, \$10, \$50, or \$100 (between–participants). In this study, the authors found that participants inferred the probabilities to be inversely related to the magnitude of the payoff, and participants' estimates influenced whether they would play the gamble or not.

These results are where the current dissertation picks up. The questions of the work presented here revolve around the *cognitive* aspects and consequences of risk–reward structures. For instance, how sensitive are people to the structure of the environment they are in, and how do they pick up these structures in the first place? And do people adapt to different risk–reward structures when making predictions and decisions under uncertainty? If risks and rewards are uncorrelated, simply assuming that high rewards are unlikely would damage functional validity. Moreover, it is unclear how much people generalize their assumptions of risk–reward relationships in different domains. Outside the lab, people often "gamble" in richer environments. For example, when betting on the outcome of a sporting event (e.g. a soccer match), people may gauge their chances of winning by relying on the prior knowledge they have about these events—but they are also given a payoff within these bets, and a pay–to–play fee. Does the payoff still influence probability estimates in such cases? In other words, how strong is the influence of the intercue relationship between payoffs and probabilities when probabilities are also tied to other cues? Before turning to these questions in more detail, I will summarize another key motivation for studying how people respond to risk–reward relationships.

Representative experimental design

Consider the choice between an 80% chance of 4000 pounds (option A) or a 100% chance of 3000 pounds (option B). Kahneman and Tversky (1979) used this gambling problem to show that people prefer the smaller, sure outcome in option B to the larger, probabilistic outcome in option A, despite this alternative offering a lower expected value. According to Kahneman and Tversky (1979) the "outcomes [used in the choice problem] refer to Israeli currency. To appreciate the significance of the amounts involved, note that the median net monthly income for a family is about 3,000 Israeli pounds" (p. 264). Also in Bernoulli's example, in which the very poor fellow gets a lottery ticket that offers a ".5 chance of winning 20,000 ducates, otherwise nothing", the lottery ticket just emerges as a windfall. In what has become known as the Allais Paradox (presented in detail in Chapter 4), Maurice Allais asked people to choose between the following sets of options (preferences denoted :

$$\text{Option A}: \quad \text{100\% chance of winning 100 million Fr.}$$

$$\text{Option B}: \begin{cases} \text{10\% chance of winning 500 million Fr.} \\ \text{89\% chance of winning 100 million Fr.} \\ \text{1\% chance of winning 0} \end{cases}$$

$$\text{Option C}: \begin{cases} \text{11\% chance of winning 100 million Fr.} \\ \text{89\% chance of winning 0} \end{cases}$$

$$\text{Option D*}: \begin{cases} \text{10\% chance of winning 500 million Fr.} \\ \text{90\% chance of winning 0} \end{cases}$$

It is fair to say that most people do not get to make these choices very often. In Brunswik's terms, they are experimental manipulations that could be "more like a mere homunculus of the laboratory out in the blank" (1955, p. 204). This is a direct consequence of ignoring any assumptions about links between the environment and the mind (Figure 2). If choices are considered in isolation, they are unlikely to generalize to nonlaboratory environments where many of these links between cues typically exist (Dhami et al., 2004; Brunswik, 1955).

To study these organism–environment relations appropriately, stimuli should be sampled from the decision maker's natural environment. They should be representative of the population of stimuli to which it has adapted and to which researchers wish to generalize their findings. In monetary lotteries, risks and rewards, or payoffs and probabilities, are often factorially combined. As Dhami et al. (2004) suggested, the range of variables is often arbitrary. In the case of Allais, 100 million Fr. with $p = 1.0$ is just as likely to be part of the choice set as 100 million Fr. with $p = 0.1$, and the natural correlation among variables (intercue relations in Figure 2) is eliminated. If we compare this to the systematic inverse relationship between risks and rewards in many natural domains, it almost seems that they are artificially "untied" (Brunswik, 1955). As the Allais paradox shows, monetary lotteries can even be unrepresentative if they are composed of locally non–dominated options, i.e., possess a local risk–reward relationship (e.g., gamble A offers a higher payoff x, but gamble B offers a higher probability p: $x_A > x_B$ and $p_A < p_B$).

Thus, one may question what can be learned from the confined world of monetary lotteries that Savage called "small worlds" (1954)? One way to broaden out the conclusions of monetary lotteries to a larger class of choice situations outside the lab is to make them more representative. That is, theories of risky choice and the processing thereof should also be tested in gamble environments in which higher rewards are less likely than smaller rewards within the entire set of gambles they are drawn from, especially if the high rewards are excessively large as in the case of the Allais paradox. While we do encounter large (possible) rewards in the world, for instance when learning about current lottery jackpots, we typically only encounter them together with a low chance of obtaining them.

Studying decision making in a representative structure that models the structure of the environment to which the mind is attuned (Anderson and Schooler, 1991; Todd and Gigerenzer, 2007; Todd et al., 2012) can help formulate better theories of how people make decisions outside the lab. Most of the experiments in this dissertation pit choice behavior in more representative environments—in which the inverse relationship between risks and rewards is maintained—against choice behavior in potentially less representative environments in which there is no relationship between risks and rewards. The term "less representative" for the uncorrelated case is chosen on purpose, as the relationship between risks and rewards can vary across environments. For example, in newly forming environments, in environments with little competition, in environments in which resources or rewards are not sufficiently limiting (e.g. air), or in heterogenous environments in which the high rewards are not easily detected, risks and rewards are less correlated or uncorrelated. While theories derived from laboratory studies with uncorrelated risks and rewards apply to these cases, they may only capture a fraction of how people make decisions under risk. Expanding the laboratory choice worlds to more representative worlds in which high rewards are unlikely seems a good compromise between the usual trade–off researchers face—the one between internal and external validity of their conclusions (Dhami et al., 2004).

Overview of the dissertation

In this work, I show how the mind learns, adapts to, and exploits the relationship between risks and rewards. In chapters 2, 3 and 4, these questions are investigated experimentally, using monetary gambles in which the majority of gambles our participants experience and evaluate are drawn from environments with a consistent risk–reward structure. Across experiments and conditions, risk–reward structures varied. In negative risk–reward environments, the structure was fairly representative of the many nonlaboratory environments in which risks and rewards are inversely related. In positive risk–reward environments, the structure was more blissful than can be expected from environments outside the lab, with probabilities increasing as payoffs increase. In uncorrelated risk–reward environments, risks and rewards are randomly paired as is frequently done in risky choice experiments (Pleskac and Hertwig, 2014). This last case thereby provides a good baseline environment to compare against more representative environments. Each chapter is or has been prepared for publication, and can thus also be read self–contained.[6]

Chapter 2 addresses two questions. First, how can people learn about risk–reward structures? Typically, people do not have the luxury to learn from explicit feedback, nor do they have an explicit goal to learn. Across three experiments, we showed that people seem to be good automatic processors of risk–reward structures as they go about evaluating the "goodness" of the options. The second question pertains to how learning about different risk–reward structures affects peoples decisions under uncertainty (Figure 1C, iii). An adaptive view of cognition implies that people should be willing to flexibly harness the structure as the ecological regularity varies across environments (Todd and Gigerenzer, 2007). We showed that indeed, people only estimated high rewards to be unlikely if they had learned that such a link exists. When there was no systematic relationship between risks and rewards, people withheld from estimating a high payoff to be unlikely. Moreover, decisions under uncertainty were consistent with inferred probabilities, which resulted in environment–dependent preferences. In more general terms, learning about risk–reward structures helped participants to reduce systemic uncertainty (Figure 1A), and the fact that participants *flexibly* adapted to risk–reward structures showed that the mind is aware and ready to adapt to changing environments (Figure 1B).

Chapters 3 and 4 address the question of how risk–reward structures influence decision making under risk (Figure 1C, ii). As the brief historic overview showed, many key findings in the decision sciences are based on the study of monetary gambles, and many existing theories have been discarded or modified due to evidence from gamble studies that a previous theory could not accommodate. Pleskac and Hertwig (2014) have shown that risks and rewards in nearly all empirical studies of risky choice are globally—across the all gambles in a given study—uncorrelated. As mentioned before, such types of gambles may be fairly unrepresentative of the gambles people choose among in non-laboratory environments. To what extent is this a problem for empirical studies on risky choice? Chapter 3 experimentally manipulates risk–reward structures as the context in which such options are presented. Specifically, we show that people build expectations about the structures of the options from their global choice environments. When presented with surprising options that deviate from these expectations, people slow down to evaluate them, especially

[6]This is not a cumulative, publication–based dissertation but follows it in form.

when they are “too good to be true” as Allais’ gambles would be in an environment where high rewards are typically unlikely. “Too good to be true” options can be conceptualized as positive “black swan” events in Figure 1C, iv).

Chapter 4 characterizes how risk–reward structures affect evidence accumulation in decisions under risk, across all options in the set. Briefly, we found that the expectation of uncorrelated risks and rewards triggers more rigorous processing than the expectation they are inversely related. These differences can help understanding peoples choices in newly forming markets versus satiated markets, environments with a more versus less ideal free distribution, or laboratory risky choice studies in which risks and rewards are uncorrelated versus inversely related. In all of these domains, uncorrelated environments may lead people to process the options more generously while risk–reward structures permit, and appear to promote, satisficing.

Chapter 5 takes takes an applied perspective. The starting point of this work was the observation that clinical trials offering high pay are sometimes considered ethically inappropriate, or even repugnant. One reason for this is that offering $10,000 for clinical trial participation can coerce people into participating who would not have done so otherwise. As we found, another reason for the repugnance of high–paying clinical trials is that a trial offering $10,000 is often considered to be riskier than a descriptively identical trial offering $1,000. This work extends the notion that people are aware that high payoffs are typically unlikely to the loss domain: If high payoffs are not unlikely, they are assumed to come at a cost. In other words, people may infer that high payments compensate for high risk because they assume there are no “free lunches” in clinical trial markets—just like nothing comes for free in many other choice environments. In Chapter 6, I synthesize the results from Chapters 2–5 and conclude with directions for further research.

References

Anderson, J. R. and Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 2(6):396–408.

Barron, G. and Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, 16(3):215–233.

Bernoulli, D. (1954). Exposition of a New Theory on the Measurement of Risk (1748). *Econometrica*, 22(1):23–36.

Bowles, S. (2006). Group Competition, Reproductive Leveling, and the Evolution of Human Altruism. *Science*, 314(5805):1569–1572.

Boyd, R. (2006). The Puzzle of Human Sociality. *Science*, 314(5805):1555–1556.

Brunswik, E. (1944). Distal focussing of perception: Size constancy in a representative sample of situations. *Psychological Monographs*, 56(254):1–49.

Brunswik, E. (1952). The conceptual framework of psychology. *Journal of Consulting Psychology*, 16(6):475–475.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62(3):193–217.

Camerer, C. and Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty*, 5(4):325–370.

Camerer, C. F. (2000). Prospect Theory in the Wild: Evidence from the Field. In *Choices, Values, and Frames*, pages 288–300.

Davies, N. B., Krebs, J. R., and A., W. S. (2012). *An introduction to behavioral ecology. 2nd.* Wiley-Blackwell, New York, NY.

Denrell, J. (2007). Adaptive learning and risk taking. *Psychological Review*, 114(1):177–187.

Dhami, M. K., Hertwig, R., and Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. 130(6):959–988.

Dreisig, H. (1995). Ideal Free Distributions of Nectar Foraging Bumblebees. *Oikos*, 72(2):161–172.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics*, 75(4):643–669.

Engel, C. (2011). Dictator games: a meta study. *Experimental Economics*, 14(4):583–610.

Fox, C. and Ülkümen, G. (2011). *Two dimensions of uncertainty.* Universitetsforlaget, Oslo.

Fretwell, S. D. and Lucas, H. L. J. (1970). On territorial behavior and other factors influencing habitat distribuion in birds. I. Theoretical development. *Acta Biotheoretica*, 14:16–36.

Gibson, J. (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston, MA.

Gigerenzer, G., Todd, P., and the ABC Research Group (1999). *Simple Heuristics That Make Us Smart*. Oxford University Press, New York, NY.

Goldstone, R. L. and Ashpole, B. C. (2004). Human foraging behavior in a virtual environment. *Psychonomic Bulletin & Review*, 11(3):508–514.

Goldstone, R. L., Ashpole, B. C., and Roberts, M. E. (2005). Knowledge of resources and competitors in human foraging. *Psychonomic Bulletin & Review*, 12(1):81–87.

Güth, W. (1995). On ultimatum bargaining experiments - A personal review. *Journal of Economic Behavior and Organization*, 27(3):329–344.

Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, 3(4):367–388.

Hacking, I. (1975). *Emergence of Probability*. Cambridge University Press, Bristol.

Hau, R., Pleskac, T. J., Kiefer, J., and Hertwig, R. (2008). The description-experience gap in risky choice: The role of sample size and experienced probabilities. *Journal of Behavioral Decision Making*, 21(5):493–518.

Hertwig, R. and Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences*, 13(12):517–523.

Hoffman, E., McCabe, K. A., and Smith, V. L. (1996). On expectations and the monetary stakes in ultimatum games. *International Journal of Game Theory*, 25(3):289–301.

Holzer, H. J., Katz, L. F., and Krueger, A. B. (1991). Job Queues and Wages. *The Quarterly Journal of Economics*, 106(3):739–768.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Kennedy, M. and Gray, R. D. (1993). Can Ecological Theory Predict the Distribution of Foraging Animals? A Critical Analysis of Experiments on the Ideal Free Distribution. *Oikos*, 68(1):158–166.

Knight, F. H. (1921). *Risk, Uncertainty, and Profit*. Sentry Press, New York, NY.

Kozyreva, A., Pleskac, T., Pachur, T., and Hertwig, R. Interpreting uncertainty: A brief history of not knowing. In Hertwig, R., Pachur, T., and Pleskac, T., editors, *Taming Uncertainty*. MIT Press, Boston, MA. forthcoming.

Krueger, A. B. (1988). The determinants of queues for federal jobs. *Industrial and Labor Relations Review*, 41(4):567–581.

Laplace, P. S. M. (1776). *A Philosophical Essay on Probabilities*.

Levitt, S. and List, J. (2009). Field experiments in economics: The past, the present, and the future. *European Economic Review*, 53(1):1–18.

Liu, B., Huang, Z., Xu, G., Jin, Y., Chen, Y., Li, X., Wang, Q., Song, S., and Jing, J. (2016). Altruistic sharing behavior in children: Role of theory of mind and inhibitory control. *Journal of Experimental Child Psychology*, 141:222–228.

Lopes, L. L. (1983). Some thoughts on the psychological concept of risk. *Journal of Experimental Psychology: Human Perception and Performance*, 9(1):137–144.

Luce, R. D. and Raiffa, H. (1957). *Games and Decisions*. Dover Publications, Inc., New York, NY.

Meder, B., Lec, F. L., and Osman, M. (2013). Decision making in uncertain times: What can cognitive and decision sciences say about or learn from economic crises? *Trends in Cognitive Sciences*, 17(6):257–260.

Ohashi, K. and Yahara, T. (2002). Visit larger displays but probe proportionally fewer flowers: counterintuitive behaviour of nectar-collecting bumble bees achieves an ideal free distribution. *Functional Ecology*, 16(4):492–503.

Payne, J. W., Bettman, J. R., and Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3):534–552.

Pleskac, T., Diederich, A., and Wallsten, T. (2015). Models of decision making under risk and uncertainty. In Busemeyer, J., Wang, Z., Townsend, J. T., and Eidels, A., editors, *The Oxford Handbook of Computational and Mathematical Psychology*. Oxford University Press, Oxford, UK.

Pleskac, T. J. (2008). Decision making and learning while taking sequential risks. *Journal of Experimental Psychology-Learning Memory and Cognition*, 34(1):167–185.

Pleskac, T. J., Conradt, L., Leuker, C., and Hertwig, R. (in prep). The ecology of competition: A linchpin for risk–reward structures. *Manuscript in prepartion*.

Pleskac, T. J. and Hertwig, R. (2014). Ecologically rational choice and the structure of the environment. *Journal of Experimental Psychology: General*, 143(5):2000–2019.

Rieskamp, J. and Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, 135(2):207–236.

Savage, L. J. (1954). *The Foundations of Statistics*. John Wiley & Sons, New York, NY.

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, 237(4820):1317–23.

Sokolowski, M. B., Tonneau, F., and Freixa i Baqué, E. (1999). The ideal free distribution in humans: An experimental test. *Psychonomic Bulletin & Review*, 6(1):157–161.

Stewart, N., Chater, N., and Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, 53(1):1–26.

Sutton, R. S. and Barto, A. G. (1998). Introduction to Reinforcement Learning. *Learning*, 4(1996):1–5.

Todd, P. M. and Gigerenzer, G. (2007). Environments that make us smart: Ecological rationality. *Current Directions in Psychological Science*, 16(3):167–171.

Todd, P. M., Gigerenzer, G., and the ABC Research Group (2012). *Ecological Rationality: Intelligence in the World*. Oxford University Press, New York, NY.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323.

Tymula, A., Rosenberg Belmaker, L. A., Roy, A. K., Ruderman, L., Manson, K., Glimcher, P. W., and Levy, I. (2012). Adolescents' risk-taking behavior is driven by tolerance to ambiguity. *Proceedings of the National Academy of Sciences*, 109(42):17135–17140.

Ungemach, C., Stewart, N., and Reimers, S. (2011). How incidental values from the environment affect decisions about money, risk, and delay. *Psychological Science*, 22(2):253–260.

van den Bos, W. and Hertwig, R. (2017). Adolescents display distinctive tolerance to ambiguity and to uncertainty during risky decision making. *Scientific Reports*, 7.

von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press (1980), Princeton, NJ.

Weber, E. U., Shafir, S., and Blais, A. R. (2004). Predicting risk sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychological Review*, 111(2):430–445.

2 | Exploiting risk–reward structures in decision making under uncertainty

Leuker, C., Pachur, T., Hertwig, R. & Pleskac, T.J. (2018), *Cognition*, 175, 186–200.

Abstract

People often have to make decisions under uncertainty — that is, in situations where the probabilities of obtaining a payoff are unknown or at least difficult to ascertain. One solution to this problem is to infer the probability from the magnitude of the potential payoff and thus exploit the inverse relationship between payoffs and probabilities that occurs in many domains in the environment. Here, we investigated how the mind may implement such a solution: (1) Do people learn about risk–reward relationships from the environment—and if so, how? (2) How do learned risk–reward relationships impact preferences in decision-making under uncertainty? Across three experiments (N = 352), we found that participants can learn risk–reward relationships from being exposed to choice environments with a negative, positive, or uncorrelated risk–reward relationship. They were able to learn the associations both from gambles with explicitly stated payoffs and probabilities (Experiments 1 & 2) and from gambles about epistemic events (Experiment 3). In subsequent decisions under uncertainty, participants often exploited the learned association by inferring probabilities from the magnitudes of the payoffs. This inference systematically influenced their preferences under uncertainty: Participants who had been exposed to a negative risk–reward relationship tended to prefer the uncertain option over a smaller sure option for low payoffs, but not for high payoffs. This pattern reversed in the positive condition and disappeared in the uncorrelated condition. This adaptive change in preferences is consistent with the use of the risk–reward heuristic.

 doi:10.1016/j.cognition.2018.02.019 Acknowledgments: We would like to thank Jann Wäscher and Chantal Wysocki for assistance with data collection.

Introduction

In March 2016, James Stocklas won $291 million in the Florida Powerball lottery. Most people know that winning such a huge jackpot is a pretty unlikely event. Now consider his brother, Bob Stocklas. Bob bought a ticket for the same lottery at the same time as James and won just $7 (Newsome, 2016). Most people know that winning this kind of sum is far more likely than winning the jackpot. And, of course, most people are also painfully aware that not winning anything at all is much more likely than either of these events. While this story illustrates the strange vicissitudes of fortune, for our purposes it also illustrates just how comfortable people are with estimating the probability of winning from payoff magnitudes alone. How do people "know" how to estimate the chances of winning the lottery? Why do they associate the highest payoff with the lowest probability? Here, we argue that the key to understanding how the mind generates such estimates lies not within the mind alone, but how the mind is adapted to its environmental context (Anderson, 1991; Gibson, 1979; Gigerenzer et al., 2011; Marr, 1982; Shepard, 1987; Simon, 1956; Stewart et al., 2006; Perkovic and Orquin, 2017).

Beyond the lottery, risks and rewards, or payoffs and probabilities, are linked in many choice environments. Across choice environments, probably the most frequent and recurrent link between them is an inverse relationship: The higher rewards that we desire are unlikely to be obtained (Pleskac and Hertwig, 2014). However, the strength of the relationship also varies across different domains. Monetary gambles in casinos, for instance, show a near perfect (though biased) inverse relationship between payoffs and probabilities. In other domains, such as where to submit a scientific manuscript (trading off impact factor against acceptance rate), the risk–reward relationship is less strong. Moreover, a risk–reward relationship is not always a given. For instance, no relationship between risk and reward is to be expected in newly forming markets that have not yet reached an equilibrium (Pleskac and Hertwig, 2014).

After identifying the ecological structures in which the mind usually operates, one can try to establish how the mind comes to terms with those ecological structures (Brunswik and Kamiya, 1953; Simon, 1956): Risk–reward structures can be exploited in decisions under uncertainty — where people have to choose between options whose payoffs are known but probabilities are not (Knight, 1921; Luce and Raiffa, 1957; Wakker, 2010). Pleskac and Hertwig (2014) offered participants a gamble that gave them a chance to win $\$x$ at the cost of $2, and asked them to estimate the probability of winning $\$x$. Different participants were asked to consider different magnitudes of x. As the magnitude of the potential payoff increased, the estimated probabilities of winning decreased. That is, participants inferred the probabilities to be inversely related to the magnitude of the payoff. Moreover, the estimates ultimately influenced what participants chose.

Inferring a probability from the magnitude of the potential payoff might be an adaptive solution to decision-making under uncertainty — a solution that Pleskac and Hertwig (2014) refer to as the *risk–reward heuristic*. Here, we investigate two of its requirements: First, the mind has to be sufficiently sensitive to the relationship between the key variables in an environment (Brunswik, 1955; Gigerenzer et al., 1991; Gibson, 1979; Marr, 1982; Simon, 1956; Stewart et al., 2006) or even mirror the relationship from the environment (Anderson and Schooler, 1991; Shepard, 1967, 1987). Second, people should be willing to harness the structure flexibly, as the ecological regularity varies across environments (Todd and Gigerenzer, 2007). That is, there should be a link between the estimates people give and an environments' risk–reward structure. This link also means that, for instance, people should withhold from estimating a high payoff to be unlikely if appropriate (e.g., in a newly forming market). This argument can be developed further: Payoffs and (subjective) probabilities determine the value of an option, and ultimately choice. Therefore, different risk–reward environments should not only affect the estimates themselves but also decisions under uncertainty.

Figure 1 provides an overview of the assumed relationships between risk–reward structures and choice that we take in this paper. Next, we develop our hypotheses in more detail, before reporting three experiments to test them.

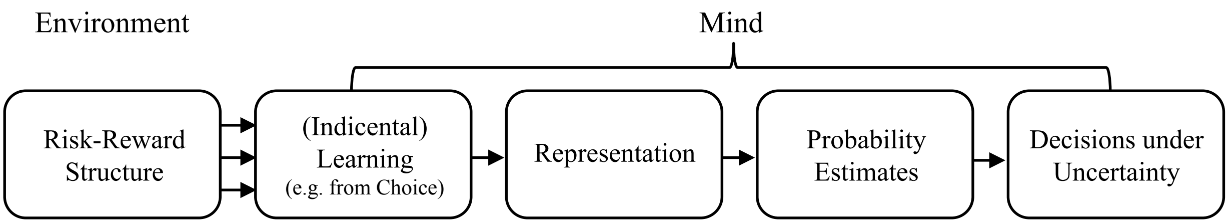


Figure 1. Summary of the assumed relationships among risk–reward structures in the world and how they ultimately shape preferences under uncertainty.

How Can People Learn Risk–Reward Structures?

In most domains, people are not explicitly told about the presence and/or direction of a risk–reward relationship. They also often do not have the luxury to learn about the relationship from explicit feedback. In this case, a risk–reward relationship would need to be acquired as people go about their primary objective when making decisions. In other words, the risk–reward relationship would seem to be learned in an unsupervised manner (without corrective feedback; Love, 2002), and incidentally (when learning is not the primary objective; Brooks, 1978; Dulany et al., 1984; Nelson, 1984; Ward and Scott, 1987; Wattenmaker, 1991; Whittlesea, 1987).[1]

Prior research suggests that via such incidental learning, people can be remarkably well attuned to statistical structures of their choice environments. For instance, they are quite good at learning the frequencies of events, even when that is not their central task (Hasher and Zacks, 1979; Hasher et al., 1987; Zacks, 2002). People also appear to encode the prices of goods and to use those prices later to evaluate the subjective worth of new values (Brown et al., 2008; Stewart et al., 2006; Olivola and Sagara, 2009; Ungemach et al., 2011), or use marginal distributions of either payoffs or probabilities in subjective evaluations thereof (Stewart et al., 2015; Walasek and Stewart, 2015). However, the risk–reward relationship is different from encoding and using (marginal) distributions of probabilities/frequencies and payoffs in that it requires people to learn a statistical regularity between probabilities and payoffs (i.e., a joint distribution). It is well known that people can learn associations between two variables (e.g. between a cue and a criterion,

[1] One might also classify this as a case of implicit learning (see, e.g., Cleeremans et al., 1998; Frensch and Rünger, 2003; Reber, 1967, 1989; Seger, 1994; Shanks and St. John, 1994). However, a typical condition for implicit learning is that individuals lack awareness of what is learned. We are thus hesitant to use this concept, as it seems that people are aware of the risk–reward relationship (Pleskac and Hertwig, 2014).

see Cooksey, 1996), and sometimes fairly quickly (Kareev, 2000; Anderson et al., 2005, but see Anderson et al., 2005). It is not known whether these findings extend to preferential choice in general; and (maybe even more importantly) to what extent people can learn that there is *no correlation* in their environment, as people may be biased to detecting structures where there are none (Olivola and Oppenheimer, 2008; Langer, 1975).

To test people's ability to learn a risk–reward relationship in an unsupervised, incidental manner, we created a learning phase in which participants encountered gambles where payoffs and probabilities were negatively correlated, positively correlated, or uncorrelated. Across experiments, we tested participants' ability to learn from different types of gambles: In Experiments 1 and 2, participants were asked to evaluate risky monetary gambles of the form "p chance of winning x, otherwise nothing." In Experiment 3, we examined to what extent participants learned different risk–reward structures from epistemic events when the probabilities were subjective (see also Tversky and Fox, 1995; Tversky and Wakker, 1995). Across experiments we also examined how different response types impacted learning with participants either choosing between gambles (Experiment 1) or indicating their willingness to sell individual gambles for (Experiments 2 and 3).

Finally, we examined in what form the risk–reward relationship is represented. In Experiments 1 and 2, we asked participants if they recognized specific gambles from the earlier learning phase. In so doing, we tested whether the risk–reward structure was learned as a "risk–reward rule" or via memory of specific gamble exemplars (Erickson and Kruschke, 1998): If it was learned via exemplars, participants should be able to recognize specific gambles from the learning phase (but not similarly structured foils).

(How) Are Different Risk–Reward Structures Exploited in Decisions Under Uncertainty?

If risk–reward structures are used in decisions under uncertainty to infer the values of missing probabilities, then this can give rise to *environment–dependent preferences*. To see this, consider an environment with a negative risk–reward relationship where high payoffs are unlikely. Someone exposed to this environment is offered a choice between an uncertain gamble with a very high payoff or a smaller, say half-as-large, certain payoff. He or she should prefer the certain payoff (i.e., the sure thing). This is because, according to the risk–reward heuristic, he or she will estimate the chances of obtaining the high uncertain payoff to be quite low and as a result the sure outcome (x) will outweigh the uncertain outcome (y) multiplied by its inferred probability ($x > p_{\text{inferred}} \times y$). The decisions of someone who has learned that risks and rewards are positively related can be expected to show the opposite pattern. Lastly, someone who has learned that risks and rewards are uncorrelated can be expected to make decisions as if the probability estimates assigned to events were independent of their payoffs. He or she may adhere to the principle of indifference, assign a probability of .5 to each outcome (Fox and Clemen, 2005; Fox and Rottenstreich, 2003), and choose the uncertain alternative equally often across payoff magnitudes.

Alternatively, preferences might be stable and simply revealed as people are asked to make decisions (i.e., revealed preference theory) (McFadden et al., 1999). A cornerstone of revealed preference theory is the principle of description invariance where the preferences and beliefs should be invariant to the description of the event (Tversky and Kahneman, 1986; Tversky et al., 1988): Given a choice between an uncertain gamble and a half-as-large certain payoff, inferred probabilities and ultimately preferences should not depend on the magnitudes of the payoffs (e.g., 1 for sure vs. 2 with an unknown probability should elicit the same preference as 1000 for sure vs. 2000 with an unknown probability). Consequently, the risk–reward environment should not impact preferences at all.

A last prediction on how risk–reward structures might impact how people deal with missing probability information in decisions under uncertainty can be derived from research on the desirability or optimism bias (Bar-Hillel and Budescu, 1995; Edwards, 1962; Irwin, 1953; Krizan and Windschitl, 2007; Sharot, 2011; Windschitl et al., 2010). Here, as payoffs become more desirable, they (or the event with which they are associated) are perceived as more likely. This prediction could either hold irrespective of the statistical relationship between risk and reward or contribute to people's environment–dependent inferences. The affect heuristic, according to which more positive overall affect towards high payoffs can mitigate perceived risk (Pachur et al., 2012; Slovic and Peters, 2006; Slovic et al., 2004), would yield a similar prediction. That is, both the optimism bias and the affect heuristic may support the belief that—probably within limits—high payoffs are by no means unlikely.

Overview of Experiments

Exp.	Learning phase		Test phase (condition-independent)	Aim of experiment
	Task	Conditions		
1	Choice	Negative Uncorrelated	Decisions under uncertainty Payoff-probability estimation Recognition	Incidental learning of risk–reward structures Influence on decisions under risk and uncertainty
2	WTS	Negative Positive Uncorrelated	Decisions under uncertainty Payoff-probability estimation Recognition Probability-payoff estimation	Incidental learning of a positive risk–reward structure Influence of type of learning phase task Influence on decisions under uncertainty
3	WTS	Negative risk Positive risk Negative uncertain Positive uncertain	Decisions under uncertainty Subjective probability estimation Payoff-probability estimation	Gambles with epistemic events Incidental learning under risk vs. uncertainty Influence on beliefs about events

Table 1. Overview of Experiments and Conditions. Learning phase stimuli were condition-dependent. Test phase tasks were condition-independent. WTS: Willingness to sell.

We conducted three experiments, each consisting of a condition-dependent learning phase and a test phase (Table 1). In Experiments 1 and 2, learning environments consisted of gambles of the form "p chance of winning x, otherwise nothing." In Experiment 3, the gambles were about an epistemic event, namely, whether the maximum temperature in Berlin on a particular day in 2011 fell within a given range. This design allowed us to examine how well participants learned risk–reward structures from choice material in which the probabilities were not explicitly stated.

In all three experiments, environments were constructed such that across the gambles probabilities and payoffs were either negatively correlated, positively correlated, or uncorrelated. Importantly, participants were neither informed about the risk–reward structures nor asked to attend to them; instead they merely experienced the structure by evaluating monetary gambles. After the learning phase, we tested how exposure to different risk–reward environments impacted participants' preferences among uncertain options. Participants then completed payoff-probability estimation tasks, which we used to test whether they had

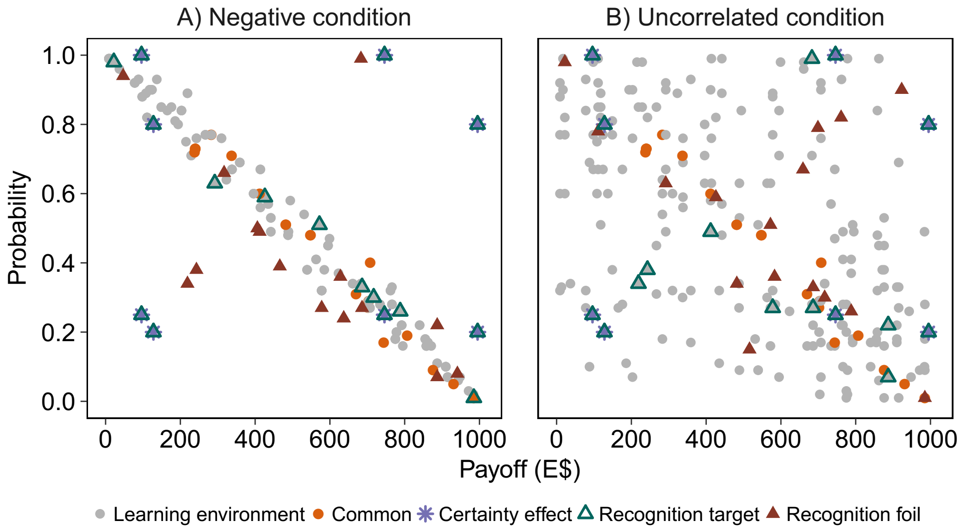


Figure 2. Stimuli used in Experiment 1. The learning phase consisted of 200 condition–dependent gambles (depicted here) that appeared in 100 nondominated gamble pairs in the experiment. Common gambles (10 pairs) and certainty-effect gambles (4 pairs) were randomly interspersed in the second half of the learning phase, allowing us to study condition-dependent changes in decisions under risk. Dominated options not depicted.

learned the risk–reward structure. In Experiments 1 and 2, they also completed a gamble recognition task that tested whether the structure was learned via memory of specific exemplars of gambles or as a rule.

Experiment 1: Do People Learn Negative vs. Uncorrelated Risk–Reward Environments and Exploit Them in Decisions Under Uncertainty?

Our first experiment had an exploratory focus. We designed it to examine how the risk–reward structure impacts decision making under both risk (probabilities given) and uncertainty (probabilities missing). To this end, we exposed participants to different risk–reward environments, asking them to choose between two nondominating gambles of the form "p chance of winning x, otherwise nothing." Between participants, the gambles were selected from one of two environments. In the negative environment, there was a negative (linear) relationship between payoffs and probabilities across all possible gambles. In the uncorrelated environment, payoffs and probabilities were randomly paired. We hypothesized that participants would learn about the risk–reward structures as a consequence of their primary task, which was to choose the alternative they preferred.

To examine how the different risk–reward structures impacted decision making under risk, about halfway through this learning phase we included gambles common to both conditions, including eight gambles designed to examine the certainty effect (Figure 2). However, we found very little differences between the conditions with respect to choices under risk. For instance, we found the certainty effect in both conditions. As our article is focused on decisions under uncertainty, these analyses on decisions under risk are reported in the Supplementary Material (see also Leuker et al., 2018, 2017).

After the learning phase, participants completed three tasks that were identical across both conditions (test phase) (see also Table 1). The first task was the decision making under uncertainty task designed to test the environment-dependent preference prediction of the risk–reward heuristic. We then tested to what extent participants learned the respective risk–reward structure by explicitly asking them to estimate probabilities when presented with new payoffs. Finally, we administered a recognition task to investigate whether participants remembered specific gambles (exemplars) or whether they had extracted a "risk–reward rule" from the learning phase.

Method

Participants

We set a target sample size of 60 participants. In total, the sample comprised 62 adults (32 females, mean age = 25.6, $SD = 3.4$, proportion students = .93) from the participant pool maintained at the Max Planck Institute for Human Development (32 in the negative condition, 30 in the uncorrelated condition).[2] All experiments were approved by the IRB of the Max Planck Institute for Human Development. Participants gave signed informed consent prior to the experiment; they were paid a fixed rate of 10 €/hour plus a bonus contingent on their choices.

Decisions under risk (learning phase)

During the learning phase, participants repeatedly chose between two monetary gambles of the form "p chance of winning x, otherwise nothing." All payoffs across all three conditions were expressed using an experimental currency, E\$. We did this with the goal of minimizing the impact of outside norms associated with specific currencies on the experiments. Each individual gamble was selected from either a negative or uncorrelated risk–reward environment (Figure 2). The experiment's negative risk–reward environment consisted of 200 gambles (but 100 nondominated gamble *pairs*) that followed a negative linear (though slightly noisy) risk–reward environment. Precise details on how the gambles were created and paired can be found in the Supplementary Materials (plus code on the OSF). For the uncorrelated condition, we took the 200 gambles (100 gamble pairs) used in the negative condition, but now randomly linked probabilities and payoffs. If any of the gamble pairs had stochastically dominated options (i.e., $p_A > p_B$ and $x_A > x_B$), we switched the probabilities of gambles A and B. We did this to maintain the marginal distributions of payoffs and probabilities across both conditions (see Stewart et al., 2006, 2015).

In both conditions, we included five dominated options that we used as 'catch trials' to identify participants who did not pay attention. In addition, 14 identical gamble pairs appeared in both conditions. Ten of these pairs were based on the procedure for the negatively correlated risk–reward environment. The other four pairs were designed to examine the certainty effect (see Supplementary Material, Table A1). Across participants, we randomized the positions of the gambles on screen, and counterbalanced the location of payoffs and probabilities (top/bottom).

Decisions under uncertainty (test phase)

We drew 20 random payoffs (range E\$1–1000) and gave them a probability of "?". Each uncertain payoff y was then matched with a half-as-large certain option (probability 100%). For example, one pair of

[2] Ten other participants also completed the experiment, but a coding error in the computerized experiment corrupted their data.

gambles asked participants to chose between a 100% chance of winning E\$50 and a "?" chance of winning E\$100. We also included 20 filler trials, in which the certain payoffs were also smaller, but created using different fractions of the uncertain payoffs. We did this to ensure that participants attended to payoffs, probabilities and uncertainty in each trial. The location of the uncertain option was counterbalanced across participants. The location of the payoffs and probabilities (top/bottom) matched the location used during the learning phase.

Payoff–probability estimation task (test phase)

To test the extent to which (individual) participants had learned about condition-dependent risk–reward relationships, we drew 10 random payoffs (range E\$1–1000), and later asked participants for their estimates of the associated probabilities.

Recognition (test phase)

Finally, to test whether participants recognized specific gambles that did not fit the risk–reward structure of a condition (incoherent gambles "off" the slope, see Figure 2), we asked participants whether they recognized (yes or no) gambles from the learning phase. The recognition task included (1) certainty-effect gambles as a particular case of exemplars that people may recall particularly well, (2) eight environment gambles from the learning phase as a subsample of exemplars that people may have encoded during learning, (3) eight environment gambles that did not appear in the learning phase (but matched the gamble structure of the condition), and (4) eight environment gambles that appeared in the other condition (thus did not match the gamble structure of the condition). This resulted in 32 cued-recognition trials (16 targets, 16 foils; see triangles in Figure 2).

Procedures

Participants were randomly assigned to either the negative or the uncorrelated condition. Participants were told that they would be asked to make a series of choices between monetary gambles in the first part of the experiment, and that there would then be some additional questions. All experiments were coded in PsychoPy (Peirce, 2007). Screenshots of all experiments can be found in the Supplementary Materials.

In the learning phase, participants saw a fixation cross (for 500 ms) before making a choice between two gambles. The chosen option was highlighted for 500 ms (by a red rectangle around the gamble), but participants did not receive any feedback about their actual payoff until the very end of the experiment. Participants took self-paced breaks after blocks of 30 trials. Gambles were presented in random order. The gambles common to both conditions were randomly interspersed after 50 condition-dependent learning trials.

To link the test phase with the learning phase, we told participants that they would see gambles that were structured similarly to the gambles they had experienced previously, and asked them to think back to these gambles when completing the task given. The order of tasks in the test phase was counterbalanced, with one constraint: Participants always completed the probability estimation task *after* the decision under uncertainty trials to minimize experimental demand effects in the choice task (i.e., prompting participants to infer probabilities from payoff magnitudes).

At the end of the experiment, we played out the chosen option of 20 randomly drawn trials of the learning phase. Bonuses (between 1.92 € and 7.74 €, with $E\$1000 = 1$ €) were added to the regular payment.

Analyses

We used a Bayesian approach to data analysis (Kruschke, 2014). Specifically, we applied Bayesian Generalized Linear Mixed Models using Stan in R for regression analyses with the rstanarm package (Stan Development Team, 2016). Unless otherwise noted, we entered participant as a grouping factor to account for individual variation beyond condition-dependent effects. Choice data were analyzed using logistic regressions; estimation data (restricted between [0,1]) were modeled after response data had been transformed to a logit scale. When plotting the posterior-predictive fits of the statistical model, we back-transformed the estimates using the inverse logit. When analyzing probability estimates, we analyzed both the estimates and the logit transformations of the estimates. For ease of interpretation, we report the results from the analyses using untransformed estimates (both analyses resulted in qualitatively identical conclusions).

We ran three chains using a Markov Chain Monte Carlo sampler to draw from posterior distributions of parameters. Depending on model complexity, we ran 10,000–30,000 samples per chain (to ensure an effective sample size of > 10,000 for each regressor) and set a burn-in of 500 samples. We investigated (convergence of) our posteriors through visual inspection and the Gelman–Rubin statistic (Gelman and Rubin, 1992). In general, we report the mean of the posterior distribution of the parameter or statistic of interest and two-sided 95% equal tail credible intervals (CI) around each value. Our focus is on estimating the effects of particular conditions and our analyses reflect this goal; in comparing the conditions, however, the crucial issue was whether the credible values included 0 or not.

Results

Decisions under risk (learning phase)

We examined choices in the learning phase to see how different risk–reward environments impacted decision making under risk. Four participants chose a stochastically dominated option once, all in the negative condition. The differences in expected values in these trials were small ($EV_{\mathrm{abs}} = E\$60$ and $E\$5, EV_{\%} = 6.0\%$ and 0.5%) and thus potentially hard to detect. We therefore included these participants' data in further analyses.

Choices between gambles were consistent with standard theories of choice: Participants chose the higher expected value gamble in 79% of all trials ($OR = 5.38$, $b = 1.68$, $CI = [1.48, 1.89]$), and this preference did not differ between the environments ($OR = .82$, $b_{\mathrm{negative}} = -0.19$, $CI = [-0.68, 0.30]$). Moreover, larger EV differences leading to more EV-maximizing choices ($b_{\mathrm{EV}} = .008$, $CI = [.007, .009]$, in a logistic regression with EV differences, higher EV, and condition as predictors). This pattern of results persisted when we compared choices in the subset of common gambles only. In addition, there were no differences in participants' subjective evaluations of payoffs and probabilities as modeled by prospect theory (Tversky

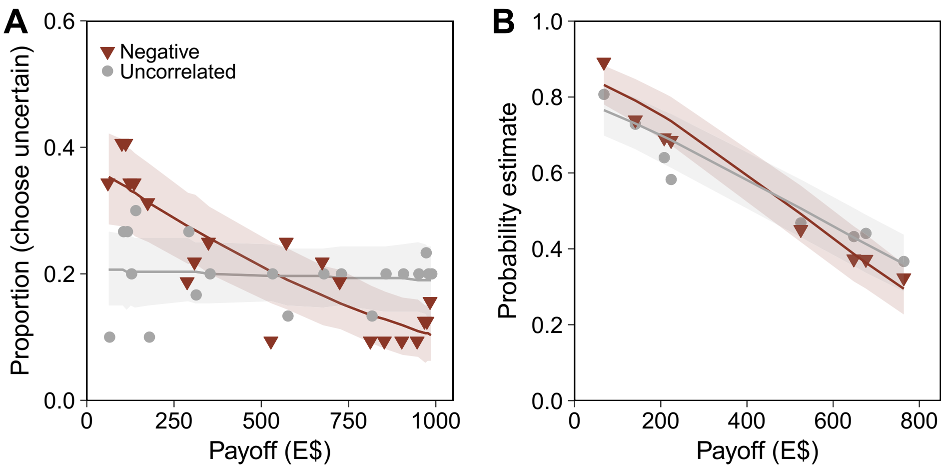


Figure 3. (A) Proportion of times the uncertain option was chosen in the decisions under uncertainty task. Participants in the negative, but not the uncorrelated, condition chose the gamble more for low and less for high payoffs. (B) Average estimated probabilities for each of the possible payoff levels in the payoff–probability estimation task. Participants in both conditions estimated an overall negative risk–reward relationship. The solid lines are the posterior predicted means from the respective regression and the ribbons reflect the 95% posterior predictive distribution.

and Kahneman, 1992) (see Supplementary Material for further details and other hypotheses we tested). In sum, we did not find evidence that manipulated risk–reward structures systematically impacted decision-making during the learning phase.

Decisions under uncertainty (test phase)

Did payoff levels shape preferences depending on the risk–reward structure experienced? In decisions under uncertainty, participants in both conditions preferred the sure option over the uncertain one ($M_{\text{uncertain}} =$ .21, $b = 1.89$, $CI = [1.44, 2.40]$). However, as predicted, the strength of preference depended on the learned risk–reward environment and the payoff magnitude offered in the gambles (Figure 3A). Specifically, participants in the negative condition chose the gamble more for low payoffs and less for high payoffs, in contrast to the uncorrelated condition ($b = 1.99$, $CI = [1.03, 2.97]$, payoff × condition interaction). These choices are consistent with participants in the negative condition inferring probabilities from payoffs, based on the risk–reward structure experienced. In the uncorrelated condition, participants tended to choose the sure thing irrespective of payoff magnitude ($M_{\text{sure}} = .19$, gray line in Figure 3A).

Payoff–probability estimation task (test phase)

Did inferred probabilities reflect previously learned risk–reward structures? Figure 3B shows participants' estimates of the probability of winning a range of payoffs. A negative risk–reward relationship was observed in both conditions ($b_{\text{negative}} = -.78$, $CI = [-.84, -.72]$, $b_{\text{uncorrelated}} = -0.57$, $CI = [-0.63, -0.51]$), but it was stronger in the negative condition ($b = -0.22$, $CI = [-0.30, -.13]$, condition × payoff interaction; in a regression with condition, payoff, and condition × interaction as predictors, using a normal link function). The results in the uncorrelated condition were unexpected in that the choice data suggest that participants had a strong prior expectation that they were not in an environment with a negative risk–reward structure.

To what degree did the learned relationship predict choices in the uncertainty task at the individual level? To investigate this, we first obtained a (risk–reward) slope for each participant through a random participant term when regressing probability estimates onto payoff magnitudes. This slope served as a measure of participants' judged risk–reward relationship. Steeper slopes indicate a stronger decrease in probability estimates as payoffs increase. This should lead to a stronger tendency to prefer the sure thing in the decisions under uncertainty task as the payoff magnitude increased. To examine this prediction, we entered these risk–reward slopes in a regression predicting choices from the risk–reward slopes, payoff, environmental condition, and the payoff × condition interaction. The regression showed that steeper risk–reward slopes predicted a stronger tendency to choose the sure thing as payoffs increased, but only for participants in the negative condition (payoff magnitude × slope, $b = 2.58$, $CI = [0.34, 4.80]$). Individual risk–reward estimates in the uncorrelated condition were not associated with choosing the uncertain option (payoff magnitude × slope × uncorrelated interaction, $b = -2.09$, $CI = [-4.79, 0.62]$; modeled in a fixed effects logistic regression, results plotted in Supplementary Figure A13A). This result speaks against the possibility that participants in the uncorrelated condition used their subjective estimates across payoffs in decisions under uncertainty. Instead, they estimated an overall negative risk–reward relationship but were averse to uncertainty in their choices across payoffs.

Recognition (test phase)

Results from the decisions under uncertainty task imply that participants were somewhat sensitive to the negative risk–reward relationship. How did they learn that relationship? Did they memorize exemplars from the learning phase? Results from the gamble recognition task suggest that participants were overall unable to discriminate targets from foils.[3] However, participants did show a bias toward stating that they recognized specific gambles (i.e., saying "Yes"): Of the eight gambles used to study the certainty effect, four fit the negative risk–reward structure (i.e., were "off" the slope) and four did not (see Figure 2A). For gambles that were inconsistent with a negative risk–reward structure (i.e., structured as the bottom left and top right gambles in Figure 2), participants tended to indicate not having seem them previously ($M_{\text{yes}} = .28$, $b = -0.73$, $CI = [-1.25, -0.22]$). This effect was more pronounced for the negative condition ($M_{\text{yes}} = .17$, $b = -1.14$, $CI = [-1.93, -0.37]$; logistic regression using risk–reward structure, condition, and their interaction as predictors, and participant as a grouping factor).

Thus, it is unlikely that participants encoded specific exemplars from the learning phase. Instead, participants in the negative condition may have abstracted a rule that they then used to assess the degree to which the stimuli were consistent with a negative risk–reward relationship. One limitation of the results from the gamble recognition task is that the stimuli set did not include any foils mimicking the structure of the certainty-effect gambles (namely, gambles located at the margins of the payoff–probability space). Instead, all of the extreme gambles were targets. We addressed this issue in the next experiment.

[3] Modeling the data in a signal detection theory framework makes this point clear: The discriminability parameter d' was centered at 0 ($M_{\text{negative}} = 0.0$, $CI = [-0.47, 0.47]$, $M_{\text{uncorrelated}} = 0.00$, $b = 0.00$, $CI = [-0.67, 0.67]$). In addition, participants did not show any systematic response biases in either condition (criterion c, $M_{\text{negative}} = 0.00$, $CI = [-0.27, 0.27]$), $M_{\text{uncorrelated}} = 0.00$, $b = 0.00$, $CI = [-0.41, 0.39]$).

Summary

Experiment 1 exposed participants to either a negative or an uncorrelated risk–reward structure. The risk–reward structure led to environment-dependent preferences under uncertainty. In the negative risk–reward condition, participants were more likely to prefer the uncertain option with lower payoffs, and their learned risk–reward relationship explained this preference. In the uncorrelated condition, choosing the uncertain alternative was unrelated to payoff magnitudes and estimated risk–reward relationships. Finally, participants (incorrectly) reported not having seen gambles when those gambles were at odds with the negative risk–reward structure, suggesting that they had encoded the overall risk–reward structure as a rule, rather than encoding specific payoff–probability exemplars.

Surprisingly, a majority of participants in the uncorrelated condition estimated an overall negative risk–reward relationship in the estimation task. Though, we should emphasize, their estimates were less extreme than the negative condition and their choices in the decision making under uncertainty task did not reflect this pattern (both at the individual and group level). Nevertheless, we offer two possible explanations for the negative risk–reward relationship in the estimates for the uncorrelated condition. First, participants in the uncorrelated condition may have an ecologically informed bias to report a (negative) risk–reward relationship even when none exists (c.f., Langer, 1975). Second, although there was no risk–reward relationship across all gambles in the uncorrelated condition, there was what might be called a local risk–reward relationship within each trial of the learning phase. As participants chose between stochastically nondominated options in the learning phase, gamble A will always have a higher payoff but lower probability than gamble B, or vice versa. Thus, participants may have learned a risk–reward relationship from the local as opposed to the global risk–reward relationship. This was similar in the uncertainty task (the uncertain option was always larger than the sure thing). In Experiment 2, we modified the learning phase so that a local risk–reward relationship was not present.

Experiment 2: Do People Learn and Exploit a Positive Risk–Reward Relationship?

Experiment 2 sought to replicate and extend the finding that participants are sensitive to risk–reward relationships and harness them in making decisions under uncertainty. To do so, we added an environment with a positive risk–reward structure creating a rather blissful structure where the larger the payoff the more likely it is to occur. Given such an idealistic structure is arguably less prevalent outside the lab, it can provide a stronger test of how well participants adapt to different risk–reward structures. The additional positive risk–reward environment affords a stronger test of how risk–reward environments may create environment-dependent preferences. Specifically, participants in the negative condition should prefer the uncertain option for low payoffs and the sure thing for high payoffs. Conversely, participants in the positive condition should prefer the sure thing for low payoffs and the uncertain option for high payoffs. Additionally, preferences for the uncertain option in the uncorrelated condition would be independent of payoff magnitudes.

To create a learning phase without a local risk–reward structure, we asked participants—instead of

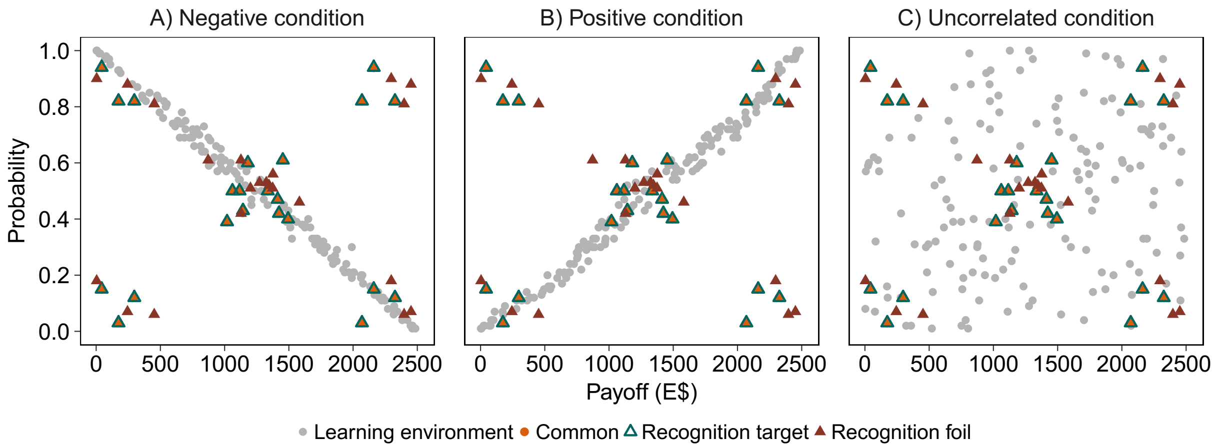


Figure 4. Stimuli used in Experiment 2. The learning phase consisted of 150 condition-dependent gambles, and 22 common gambles (triangles).

choosing between two gambles as in Experiment 1—to state the price for which they would be willing to sell (WTS) a single gamble presented at each trial for. Based on our findings from Experiment 1, we predicted that participants would learn about risk–reward relationships incidentally while pricing the gambles.

Finally, we sought to better understand the degree to which participants encoded the risk–reward relationship as a rule. We did so by modifying the gamble recognition task so that target and foil gambles were structured equally, especially at the four extremes (Figure 4). We hypothesized that participants would not distinguish between targets and foils—which would be difficult to do—but would respond based on the gambles' fit with previously experienced risk–reward structures.

Method

Participants

We recruited 90 participants (53 females, mean age 24.7, $SD = 4.1$, proportion students = .72) from the participant pool at the Max Planck Institute for Human Development. Each participant completed the experiment in exchange for a show-up fee of €10 and a performance-contingent bonus. Participants in Experiment 1 were excluded from the recruitment process.

Decisions under risk (learning phase)

The methods were largely the same as in Experiment 1; here, we summarize key differences. We used a larger payoff range (E\$1.01–2500, disclosed conversion rate E\$2500 = €1). Figure 4 depicts the three risk-reward environments used in the current experiment (150 gambles per environment with identical marginal distributions; see Supplementary Materials for details, and code on the OSF). Briefly, to create the positive risk–reward condition, we took the gambles in the negative condition and reversed the order of probabilities such that the highest probabilities were now associated with the highest payoffs and vice versa.

To study condition-dependent differences in how participants priced (identical) gambles, we included 22 gambles common to all three conditions (10 in the center, 3 at each margin; see triangles in Figure 4). In total, this procedure resulted in 172 risky gambles per risk–reward condition while controlling for the marginal distribution of payoffs and probabilities across all three conditions.

Decisions under uncertainty (test phase)

For the uncertainty task, we created gamble pairs with low ($1-250E\$$), intermediate ($1125-1375E\$$), and high payoffs ($2250-2500E\$$) (10 pairs each). As in Experiment 1, the uncertain option's payoff (probability "?") was half as big as the certain option's payoff. In a typical pair, participants chose between a 100% chance of winning $E\$50$ and a "?" chance of winning $E\$100$. We included 30 filler trials in which the certain option was created by scaling down the uncertain option by a random factor between .1 and .9.

Payoff–probability estimation task (test phase)

We increased the number of trials such that participants estimated the probabilities associated with 20 payoff magnitudes (range $E\$1$–2500). To investigate how well participants learned the bi-directional relationship between payoffs and probabilities, we also asked participants to estimate the payoff associated with a given probability at the end of the experiment (probability–payoff estimation task). We drew 20 probabilities between 0 and 1 for this task (results mirrored the results of the probability estimation task, and are reported in the Supplementary Material, Fig. A5).

Recognition (test phase)

We used the gambles common to both conditions in the learning phase as targets and an equally generated set of gambles as foils (Figure 4, red triangles). Thus, foils were 10 gambles at the center and 12 gambles at the margins of the payoff–probability distribution space (novel random draws based on the recognition gambles procedure). This broader set (relative to Experiment 1) of 22 targets and 22 equally structured foils was used to test whether the risk–reward relationship was learned via exemplars or a rule: If participants had learned the relationship as a rule, they should indicate not having seen gambles that did not fit condition-dependent risk–reward structure (and indicate having seen gambles that did), irrespective of whether those gambles were targets or foils.

Procedure

During the learning phase, participants indicated their WTS for one gamble at a time. They took self-paced breaks after each of five blocks. Common gambles were randomly interspersed after 100 condition-dependent trials. The task was presented as a game show called "Keep or Sell?" ("Behalten oder Verkaufen?"). To motivate participants to indicate their true valuations of a gamble, we implemented a Becker-DeGroot-Marschak auction (Becker et al., 1964). The rules were as follows. Participants owned the right to play each gamble, which they could sell to the experimenter at a price they determined themselves. Prices were entered with a mouse click on a rating scale ($E\$0--2500$) and confirmed with a click on the value.

To incentivize the task, we informed participants that 10 gambles would be randomly selected and played out at the end of the experiment. The experimenter then offered a (computer-generated) buying price between 0 and the maximum payoff from the gamble. If the experimenter's price exceeded the participant's selling price, the participant sold the gamble and earned the buying price. If the participant's selling price exceeded the experimenter's buying price, the gamble was played out (e.g., 50% chance of E\$380). The dominant strategy in this task is to price a gamble based on its subjective value: Higher prices can prevent participants from selling unattractive gambles; lower prices can lead to them selling attractive gambles under value. In other words, the prices should approximate participants' subjective certainty equivalents for the gambles.

Participants completed five practice trials to ensure their proper understanding of the WTS measure. If they indicated a selling price that exceeded the maximum payoff from that gamble, participants would see a screen reminding them that (i) they would only receive counteroffers between 0 and the maximum amount to be gained in the gamble, (ii) setting an accurate price would increase the likelihood of good counteroffers, and (iii) good counteroffers would maximize the bonus to be gained from the task. After this feedback, participants set a new price for the same gamble. If they had no more questions, they proceeded to the main part of the task, in which there was no feedback.

The test phase in Experiment 2 was equivalent to that in Experiment 1 (see Table 1 for an overview of the tasks and task order). The exception was that the decisions under uncertainty task was now incentivized. In particular, five choices from the uncertainty task were randomly selected and played out. If the uncertain option was chosen then the condition-dependent probabilities were used to determine the probability of the outcome. Participants were instructed about the incentivization scheme at the beginning of the task. At the end of the experiment, we played out the randomly drawn trials from the learning phase and the uncertainty task. Bonuses (between 1.99 € and 7.82 €, with $E\$2500 = 1$ €) were added to the regular payment.

Results

Decisions under risk (learning phase)

Across all gambles, including the environment gambles, prices were strongly related to the gambles' expected values (indicated by a credible payoff × probability interaction, $b = 0.88$, $CI = [0.85, 0.91]$). Prices in the positive condition deviated slightly more from expected values compared to in the other two conditions (payoff × probability × positive condition, $b = -.07$, $CI = [-.12, -.03]$). However, these differences did not persist when we modeled certainty equivalents given for the subset of gambles common to all conditions (thereby controlling for condition-dependent stimuli features). In addition, there were no differences in participants' subjective evaluations of payoffs and probabilities as modeled by prospect theory (Tversky and Kahneman, 1992) (see Supplementary Material). In sum, and consistent with Experiment 1, participants seemed to evaluate risky gambles in a similar manner across conditions.

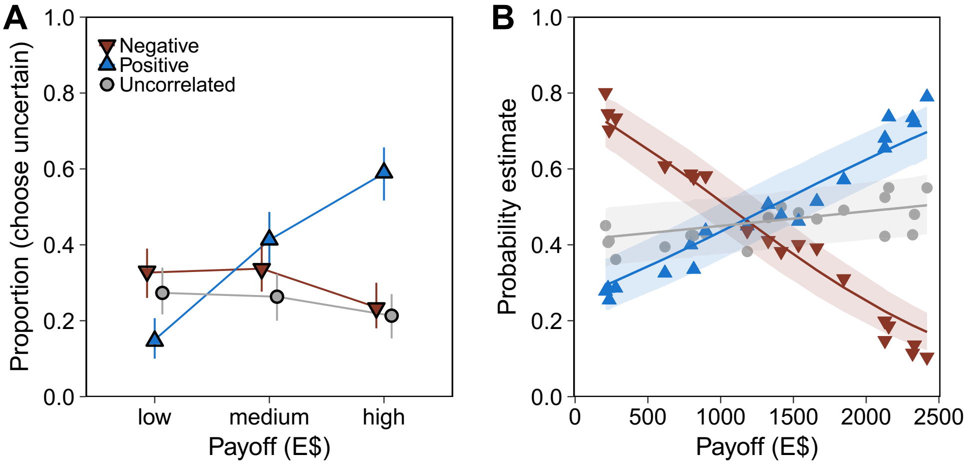


Figure 5. (A) Proportion of times the uncertain option was chosen in the decisions under uncertainty task. Participants in the positive condition chose the gamble more for high and less for low payoffs. Participants in the negative and uncorrelated conditions showed risk-averse behavior, with a low overall proportion of gamble choices. Error bars reflect the 95% posterior predictive distribution; black triangles reflect the mean of the posterior predictive distribution. Note our analysis of choices in the uncertainty task treated payoff as a continuous variable, but we binned this variable as low/medium/high for clarity in plotting the results. (B) Average estimated probabilities for each of the possible payoff levels in the probability estimation task. Participants' estimates reflected the risk–reward structures to which they had previously been exposed. The line reflects the mean of the posterior predictive distribution from the linear regression; ribbons reflect the 95% posterior predictive distribution.

Decisions under uncertainty (test phase)

How did the experienced risk–reward structures shape participants' preferences under uncertainty? Figure 5A displays the proportion of choices of the uncertain option as a function of the possible payoff level, separately for the three conditions. In general, participants were more likely to choose the certain but smaller payoff option over the uncertain option that offered a larger payoff ($M_{\text{sure}} = .63$, $b = 1.97$, $CI = [1.43, 2.55]$). However, this preference depended on the risk–reward environment to which participants had previously been exposed and on the payoff magnitude offered in the gambles. Consistent with our prediction of environment-dependent preferences, the higher the payoffs, the more often participants in the positive condition chose the gamble ($b_{\text{positive}} = 3.04$, $CI = [2.41, 3.69]$, condition × payoff interaction). When payoffs were high, participants in the positive condition chose the uncertain option in as many as 59% of trials.

As Figure 5A shows, the pattern of results was very different for participants in the negative condition, who chose the gamble slightly more for smaller payoffs and less for larger payoffs. Nevertheless, unlike the results of Experiment 1, the effects in the negative condition were rather small, and the choices were not credibly different from those in the uncorrelated condition ($b_{\text{negative}} = -.22$, $CI = [-.84, .41]$; all effects modeled in a logistic regression with the uncorrelated condition as baseline). This finding was unexpected: If participants had relied on the learning phase and exclusively used the knowledge they expressed in their probability estimates in the choice task, they should have been much more risk seeking for low payoffs, which they would have learned to be associated with high probabilities. As we will show shortly, this may be due to (or linked back to) individual variability in the learned risk–reward relationship.[4]

[4]The choice patterns for the filler gambles were identical to these results (positive condition choosing gambles more as payoffs increase; negative condition similar to uncorrelated condition); as expected, choices here also largely depended on the

Payoff–probability estimation task (test phase)

Did participants' estimates reflect risk–reward environments from the learning phase? As Figure 5B shows, the probabilities that participants estimated varied as a function of the possible payoffs. Participants in the uncorrelated condition estimated a weak positive relationship ($b = 0.11$, $CI = [0.06, 0.16]$). Estimates from participants in the other two conditions reflected the specific risk–reward structure to which these participants had previously been exposed. In the negative condition, the estimates showed a negative relationship between payoffs and probabilities ($b = -0.75$, $CI = [-0.82, -0.68]$). In the positive condition, the reverse applied but the slope was much shallower ($b = 0.35$, $CI = [0.28, 0.42]$, in a normal link regression using condition, payoff, and condition × payoff as predictors). The shallower slope was partially driven by two participants in the positive condition who estimated a negative risk–reward relationship, perhaps indicative of a possible negative risk–reward relationship 'default'/prior.

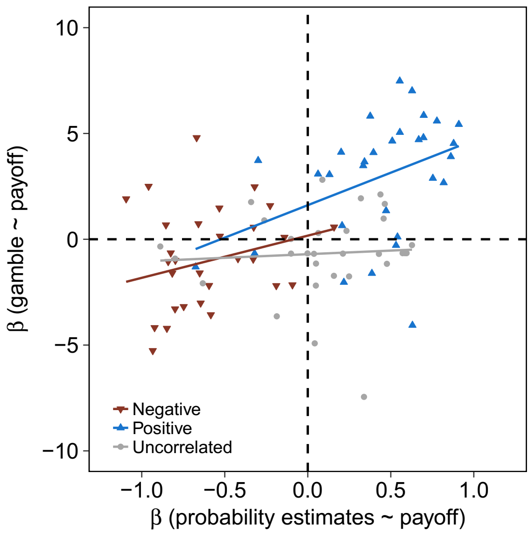


Figure 6. Individual variation in choice of the uncertain alternative (y-axis) based on estimated payoff–probability relationships (x-axis). Each data point depicts one participant. The participant-level β was estimated from a Bayesian regression with a random participant intercept for (estimates ~ payoff) and (choice ~ payoff), respectively. Values above 0 indicate higher probability estimates with increasing payoffs (x-axis), and a preference for the uncertain option with increasing payoffs (y-axis). Payoff–probability estimates in the negative and positive condition, but not in the uncorrelated condition, predicted choice.

As in Experiment 1, we also investigated the extent to which an individual's probability estimates predicted his or her choices in decisions under uncertainty. To do so, we again obtained a (risk–reward) slope for each participant through a random participant term when regressing probability estimates onto payoff magnitudes. As Figure 6 shows, the majority of slopes (plotted on the x-axis) reflected the condition to which participants had been exposed: the negative condition's slopes fell in the negative range; the positive condition's slopes fell in the positive range.

We then used the individual slopes to predict choosing the uncertain over the certain option across different payoff magnitudes. The risk–reward slopes predicted payoff-dependent preferences for the uncertain option in the two correlated conditions ($b_{\text{negative}} = 2.01$, $CI = [0.27, 3.76]$; $b_{\text{positive}} = 2.74$, $CI = [1.03, 4.45]$, slope × payoff × condition interaction in a fixed effects model using the uncorrelated condition as baseline). That is, following up on the degree to which the risk–reward conditions impacted preferences, once individual variability in the learned risk–reward relationship is accounted for, preferences in both the negative and positive condition were credibly different (and in the predicted direction) from the uncorrelated condition. Nevertheless, the link between estimates and choices was weaker in the negative condition (red vs. blue slope in Figure 6). As in Experiment 1, risk–reward slopes in the uncorrelated condition did not predict the choice of an uncertain option ($b_{\text{uncorrelated}} = .44$, $CI = [-.56, 1.43]$).

difference between certain and uncertain payoffs ($b_{\text{uncertain/certain}} = -8.84$, $CI = [-9.56, -8.16]$).

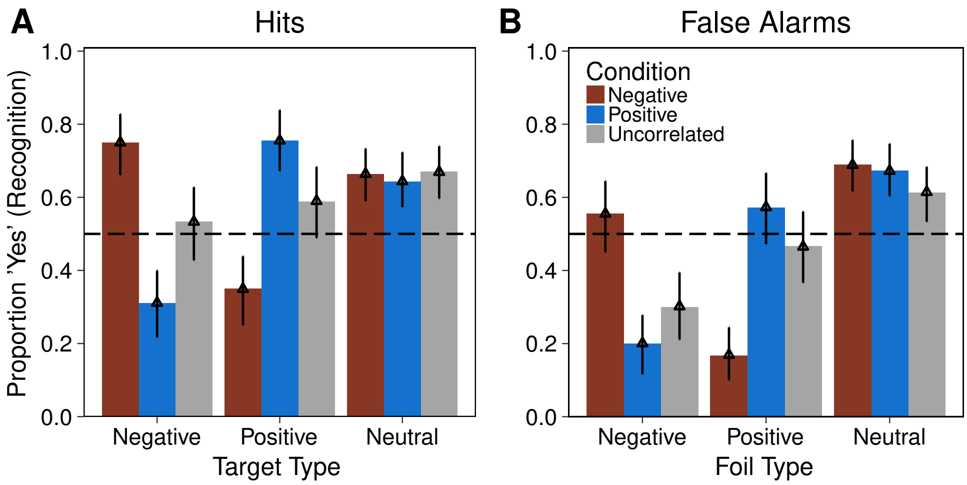


Figure 7. Proportion responding 'yes' to items in the recognition task, by stimuli characteristics and condition. Overall discriminability between targets and foils was low (similar response patterns in panels A and B). Responses depended on whether gambles fit a condition-dependent risk–reward structure. Error bars reflect the 95% posterior predictive distribution; triangles reflect the mean of the posterior predictive distribution.

Recognition (test phase)

Results from the decisions under uncertainty and estimation tasks suggested that participants learned risk–reward structures in the first phase of the experiment. But how did they represent the different structures? A comparison of panels A and B in Figure 7 shows that participants responded similarly when the gambles presented were targets versus foils, implying that they could not discriminate between them.[5]

When we broke responses down by whether or not gambles fit a condition's risk–reward structure (Figure 7), the results resembled those of Experiment 1: If gambles did not fit a condition-dependent risk–reward structure, participants indicated that they had not seen them previously, irrespective of whether these gambles were targets or foils. That is, a majority of participants in the negative condition reported not having seen gambles that were consistent with a positive risk–reward relationship ($M_{\text{yes}} = .26$, $b = -1.40$, $CI = [-1.80, -1.00]$). Conversely, a majority of participants in the positive condition reported not having seen gambles that were consistent with a negative risk–reward relationship ($M_{\text{yes}} = .26$, $b = -.85$, $CI = [-1.25, -.46]$).

What is more, participants were also more likely to report having seen gambles merely because their structure followed a condition's risk–reward structure, again irrespective of whether these gambles were targets or foils (see Figure 7). That is, participants in the negative condition were likely to report having seen a gamble if the gamble was consistent with a negative risk–reward relationship ($M_{\text{yes}} = .65$, $b = .84$, $CI = [.45, 1.23]$), and participants in the positive condition were likely to report having seen a gamble if the gamble was consistent with a positive risk–reward relationship ($M_{\text{yes}} = .66$, $b = .52$, $CI = [.13, .90]$; all results from a logistic regression using condition × stimulus type as a predictor, neutral gambles in the uncorrelated condition as baseline). Responses to neutral gambles were identical across conditions. These

[5] A signal detection analysis showed that participants did not discriminate between old and new gambles across all three conditions ($b/d'_{\text{uncorrelated}} = 0.21$, $CI = [-0.17, 0.60]$; $d'_{\text{positive}} = -0.16$, $b = -0.37$, $CI = [-0.89, 0.15]$; $d'_{\text{negative}} = -0.05$, $b = -0.26$, $CI = [-0.78, 0.27]$). There were weak, but not credible, biases towards saying 'yes' in the correlated conditions ($b/c_{\text{uncorrelated}} = -0.10$, $CI = [-0.30, 0.08]$), $c_{\text{positive}} = 0.08$, $b = 0.18$, $CI = [-0.09, 0.45]$, $c_{\text{negative}} = 0.03$, $b = 0.13$, $CI = [-0.14, 0.40]$).

gambles were consistent with all risk–reward structures, which might explain why people were equally likely to indicate that they had previously seen them (in all three conditions; $M_{\text{yes}} = .68$, bars on the right in Figure 7A and B).

Summary

Experiment 2 substantiated the findings from Experiment 1 that participants could learn risk–reward structures in an unsupervised, incidental fashion, and that they subsequently often exploited the relationship to make decisions under uncertainty. In particular, we showed that participants learned and used a positive risk–reward relationship, although this structure stands in stark contrast to the negative risk–reward relationship present in many real-world environments. Moreover, in contrast to Experiment 1, where probability estimates in the uncorrelated condition showed a negative association with payoff levels, probability estimates were now independent of payoff levels. Comparing Experiment 2 to Experiment 1 suggests that this difference is due to the learning phase, in which participants now evaluated one gamble at a time, removing any 'local' risk–reward structure naturally built into a choice task with nondominated gambles. In a payoff–probability estimation task participants used risk–reward relationships to infer probabilities from payoffs—and that the risk–reward relationship learned from pricing gambles dictated the direction of the estimates. Finally, we found further evidence of environment-dependent preferences in decisions under uncertainty. One qualification to this result is that participants in the negative condition were not as keen on choosing the uncertain option for low payoffs as we had expected (despite estimating high probabilities for these payoffs). Finally, the recognition task in Experiment 2 provided further evidence that the risk–reward relationship from the learning phase was represented as a rule rather than in terms of memorized gambles.

Across Experiments 1 and 2, the gambles used in the learning phase presented risks in terms of explicit, single numbers. Outside the laboratory, in contrast, many gambles are about epistemic events, for instance when betting on the outcome of a sporting event (e.g. a soccer match). To gauge their chances of winning for such gambles, people may rely on the prior knowledge they have about these events. In our final study, we asked how well our results generalized to choices about these epistemic events.

Experiment 3: Do People Learn About and Exploit Risk–Reward Structures from Gambles About Epistemic Events?

In Experiment 3, we examined whether people learn risk–reward relationships when the chances of winning depend on events about which they have some prior knowledge. Specifically, in both the learning and test phase, we used gambles in which winning was tied to the maximum temperature in Berlin on a particular day in 2011 falling within a certain range (e.g., "You win *E*$500 if the temperature on August 20th, 2011, was between 16 and 25°C"). We adapted a procedure from Tversky and Fox (1995) to create different events using different widths and locations of temperature ranges, so that participants should a priori have different subjective probabilities of the events occurring (see also Tversky and Wakker, 1995). As in Experiment 2, participants were asked to state prices for gambles in the learning phase. To create different risk–reward relationships, we determined the probability that a given interval would contain the maximum temperature based on the width of the interval and its proximity to the mean August temperature. We refer to these probabilities as *historical frequencies*. We paired them with payoffs between E\$1.01 and E\$2500 to create either a positive or a negative risk–reward relationship. Using these two conditions, we aimed to extend our finding that participants learn risk–reward relationships incidentally from simple monetary gambles to gambles with epistemic events.

Learning about risk–reward relationships from implied subjective beliefs alone may be challenging. Moreover, in some situations, probability estimates about epistemic events are available, such as when an meteorologist shares her belief that an event will occur. Thus, we further differentiated the learning environments, with half the participants being shown only the temperature range of the event but no explicit probability information ('learning under uncertainty') and half additionally being shown the historical frequencies ('learning under risk'). By comparing these two sets of conditions, we tested whether explicit probability information is necessary to learn the risk–reward relationship.

In the test phase, we assessed the influence of the risk–reward environments somewhat differently compared to the first two Experiments. Because the chances of the maximum temperature in Berlin falling within a particular temperature range could be inferred from the range itself, we tested for the effect of the risk–reward environment on probability estimates and choices above and beyond the information provided by the temperature range (historical frequencies). How can environment-dependent preferences emerge here? In the negative risk–reward environment that the proportion of choices of the uncertain option should increase for lower payoffs, and decrease for higher payoffs—and vice versa for the positive risk–reward environment. A similar pattern should emerge for probability estimates: When relying on a negative risk–reward structure, probability estimates for an event associated with a low payoff should be larger than probability estimates for an (otherwise comparable) event associated with a high payoff. When relying on a positive risk–reward structure, the opposite should happen.

Lastly, we examined whether participants generalized learned risk–reward relationships to other contexts. To this end, we added tasks in which participants made decisions under uncertainty and estimated probabilities about the maximum temperature falling in a particular range in Dushanbe, Tajikistan. Importantly, our results showed that participants did not rely on risk–reward structures in a context they had not been exposed to (for details see the Supplementary Material).

Method

Participants

We recruited 200 participants from the participant pool at the Max Planck Institute for Human Development in Berlin (125 females, mean age = 24.45, $SD = 4.3$, proportion students = .84) to take part in the experiment for a 10 € show-up fee and a performance-contingent bonus. Participants in Experiment 1 and 2 were excluded from the recruitment process. Due to the change in design, we expected smaller effect sizes and therefore increased our sample from 30 to 50 participants per condition. Participants were randomly assigned to one of four conditions. Due to a computer error, the responses from two participants in the uncertainty task were not saved (leaving $N = 198$).

Decisions under uncertainty vs. risk (learning phase)

During the learning phase, participants priced gambles based on the Berlin weather. To construct these gambles, we retrieved past weather data on the mean ($M = 22.7°C$) and standard deviation ($SD = 3.2°C$) of the maximum daily temperature in August in Berlin in 2011 from *accuweather.com*. We created 155 temperature ranges of varying width and location on the temperature scale (see Supplementary Figure A6). Because the maximum temperatures were approximately normally distributed, we calculated the historical frequency to approximate the probability that the maximum temperature on a given date would fall within the specified interval. We then constructed gambles such that there was either a positive or a negative risk–reward relationship holding the marginal distributions of payoffs and historical frequencies constant across conditions (see Supplementary Material A7, code on OSF).

In the learning under uncertainty condition, only the temperature range and the corresponding payoff was shown for each gamble (e.g., "E\$2300 if the maximum temperature was between 13 and 15°C on Aug 29th".) In the learning under risk condition, the historical (relative) frequency was added to the gamble (e.g., "E\$2300 if the maximum temperature was between 13 and 15°C on Aug 29th (p = 3%)."). Screenshots are shown in Supplementary Material A8.

Decisions under uncertainty (test phase)

We created a decisions under uncertainty task in which participants chose between an uncertain option that depended on the Berlin weather event having occurred ("E\$2000 if the maximum temperature was between 23 and 26°C on August 22nd") and a smaller, sure thing ("700E\$ for sure"). We varied the payoffs on two levels, to be either high (E\$2000 vs. E\$700 for sure) or low (E\$100 vs. E\$35 for sure). Participants completed 15 different trials about Berlin weather.

Subjective probability estimation task (test phase)

This task consisted of two parts. First, participants were asked to estimate their subjective probability (0–100%) of winning the gamble (i.e., the event occurring) in the decisions under uncertainty task *with payoff information*. Our key interest was the degree to which participants used the payoff information in their estimates. Participants were therefore shown the actual gamble (e.g., "E\$2000 if the maximum temperature in Berlin was between 23 and 26°C on August 22nd") and asked to judge the probability that they would win.

In a second part, participants indicated their subjective probability (0–100%) that the maximum temperature on a given day in August would fall in a given temperature range *without payoff information* (e.g., "likelihood the maximum temperature in Berlin was between 23 and 26°C on August 22nd"). The temperature ranges were identical to those used in the decisions under uncertainty task and in the subjective probability estimation task *with* payoff information.

Payoff–probability estimation task (test phase)

Participants were presented with 20 different payoffs and asked to think back to the gambles they had experienced in the learning phase. They then estimated the likelihood of winning these payoffs in the upcoming bonus trials. This task was used to test whether participants had picked up on the risk–reward structures in the learning phase.

Procedure

Participants were randomly assigned to one of four learning conditions (Negative Risk, Negative Uncertainty, Positive Risk, Positive Uncertainty). They evaluated gambles about the maximum temperature measured on a given day in August 2011 by indicating a WTS for each gamble. The instructions were adapted from Experiment 2. Here, participants were informed that a gamble's value was determined by the extent to which the temperatures were in line with the true temperatures on a given day, and by its possible payoff. The instructions for the risk conditions included an explanation of the historical frequency information added to the gambles, namely that the probability was based on typical August temperatures (i.e., "45% is the likelihood that a typical August day will fall in the temperature range given in the bet"). During the learning phase, participants took self-paced breaks between five blocks of 31 pricing trials each. The learning phase was incentivized such that prices from 10 randomly drawn trials were played out according to the Becker-DeGroot-Marschack auction procedure described in Experiment 2, but now the outcome of the gamble was determined by whether the event's temperature range actually contained the true maximum temperature.

The order of tasks in the test phase was identical across all participants. We randomized the positions of sure things versus gambles in the decisions under uncertainty task, as well as the position of the payoff amount in the gamble (above or below the event) on the trial level. The decisions under uncertainty task was incentivized such that five randomly selected choices were played out. At the end of the experiment, bonuses (between 1.28 € and 7.65 €, with $E\$2500 = 1$ €) were added to the regular payment.

Results

Decisions under risk vs. uncertainty (learning phase)

The prices suggested that participants traded off the payoff and the historical frequencies of events (effect of EV, defined as *historical frequency* × *payoff*: $b = .46$, $CI = [.40, .53]$). As expected, risky gambles that included information about historical frequencies (negative risk and positive risk conditions) were closer to the EVs of the bets than their uncertain counterparts ($b_{\text{negative\_uncertain}} = -.66$, $CI = [-.76, -.57]$, $b_{\text{positive\_uncertain}} = .58$, $CI = [.48, .67]$; 3-way interaction using a gamble's EV × risk–reward relationship × type of learning).

Decisions under uncertainty (test phase)

Did the experienced risk–reward relationships shape preferences under uncertainty? We expected this to be the case after "learning under risk" (as in Experiments 1 & 2), in particular, but also (though less strongly) after learning under uncertainty. We analyzed condition-dependent choices after controlling for the events' historical frequencies.[6] Figure A9 shows the results of this analysis. Overall, participants were less likely to choose the gamble over the sure outcome for high (E\$2000) than for low payoffs (E\$100) across conditions ($b_{\mathrm{E\$2000}} = -1.23$, $CI = [-1.62, -.85]$). Was this payoff effect moderated by learned risk–reward structures? Indeed, consistent with the risk–reward relationship they had experienced in the learning phase, this payoff effect was smaller for participants who had been exposed to a positive risk–reward relationship under risk (panel A). This effect was driven by participants in the positive condition choosing the gamble less often when the choice was associated with a E\$100 payoff—a payoff that had previously been associated with a low probability ($M_{\mathrm{gamble}} = -.12$, $b_{\mathrm{positive \times E\$100}} = -.57$, $CI = [-1.08, -.05]$, all results based on a mixed effects logistic regression controlling for historical frequencies, using learning type [risk vs. uncertain] × risk–reward relationship [negative vs. positive] × payoff level as predictors).

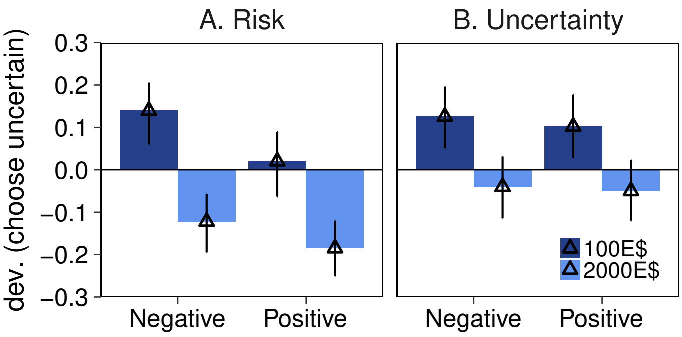


Figure 8. Decisions under uncertainty. Plots show how much participants picked the uncertain gamble after controlling for the gamble's historical frequencies. Choice proportions perfectly adjusted to the gambles' historical frequency should have a 0 deviation. Bars and triangles reflect the mean of the posterior predictive choice distribution (controlling for historical frequency); error bars indicate the 95% posterior predictive distribution (generated using historical frequencies of .5).

Learning under uncertainty did not affect choice. In sum, there is some evidence for environment-dependent preferences, namely when participants were exposed to the risk–reward relationship under risk. Participants in the positive condition became less risk seeking for low payoffs but not more risk seeking for high payoffs, as one would have expected from Experiments 1 and 2.

Subjective probability estimation tasks (test phase)

Participants were also asked to estimate the chances that a maximum temperature would fall within a given temperature range both within the context of the gamble as a whole (including payoff information associated with the event) and without this payoff information. As we were interested in how the estimates were affected by the risk–reward environments after controlling for the historical frequencies associated with the events, we report deviations from those historical frequencies.[7] Did participants rely on previously experienced risk–reward structures when gauging their chances of winning a bet about the weather? Figure A10 (A, B) shows that participants' subjective estimates were indeed guided by the payoff infor-

[6] Choices were well-adjusted to the events' historical frequencies, with an almost linear increase in the proportion of participants choosing the uncertain option as probabilities of winning based on historical frequencies increased ($b_{\mathrm{probability}} = 6.98$, $CI = [6.48, 7.51]$, see Supplementary Material S11 for posterior predictions across different historical frequencies).

[7] All participants were sensitive to historical frequencies and provided estimates that reflected these frequencies across contexts ($b = .77$, $CI = [.76, .79]$, see Supplementary Material A12).

mation. In line with our predictions and Experiments 1 and 2, in the negative conditions (panel A, left bars), subjective probability estimates were lower when temperature ranges were presented in a gamble context that offered a $E\$2000$ payoff ($b = -.10$, $CI = [-.12, -.07]$) than in a gamble context that offered a $E\$100$ payoff.

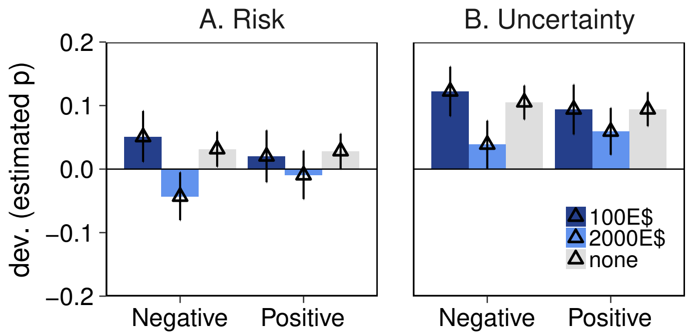


Figure 9. Subjective probability estimation tasks. Plots show deviations of participants' estimates after controlling for the gambles' historical frequencies. Estimates perfectly in line with the gambles' historical frequency should have a 0 deviation. Participants gave subjective probability estimates of winning a particular temperature bet including a payoff (blue bars) or the probability of a given temperature range alone (gray bars). Bars show posterior mean deviations from historical frequencies; error bars show 95% highest density intervals (generated using historical frequencies of .5).

This payoff effect—a difference in estimates for $E\$2000$ vs. $E\$100$, after learning under risk—was not observed in the positive condition (panel A, right bars, $b = -.04$, $CI = [-.09, .03]$ modeled in a normal link regression using using learning type [risk vs. uncertain] × risk–reward relationship [negative vs. positive] × payoff level as predictors). As Figure A10 further shows—and contrary to our predictions—the payoff effect did not flip (with higher payoffs leading to a positive deviation and lower payoffs leading to a negative deviation). A bi-product of this was that participants in the risky positive condition ended up with estimates closer to the true historical frequencies (Figure A10, panel A). For participants who had learned about risk–reward relationships under uncertainty (panel B), the between-condition effects were comparable (larger payoff effect in the negative condition, see panel B).

Did higher estimated probabilities in this task predict choices in the decisions under uncertainty task? Indeed, we found a link between estimates and choices ($b = 4.00$, $CI = [3.25, 4.78]$, main effect of estimate in a logistic regression using historical frequencies, estimates, and their interaction as predictors).

Payoff–probability estimation task (test phase)

To what extent did probability estimates reflect the experienced risk–reward structures? Figure 10 shows that, as expected, the estimates of participants in the negative condition reflected the risk–reward structure from the learning phase (slope in the negative condition = $b_{\text{negative} \times \text{payoff}} = -.40$, $CI = [-.45, -.35]$). Estimates in the positive conditions were regressive to 50% (slope = .02, $b_{\text{positive} \times \text{payoff}} = .42$, $CI = [.35, .49]$, interaction effect using the negative condition as baseline). Figure 10 (panels A vs. B) also shows that the results were identical for the risky and uncertain learning conditions ($b_{\text{uncertain}} = .02$, $CI = [-.03, .08]$, all estimates modeled in a normal link regression, using learning type (risk vs. uncertain) × risk–reward relationship (negative vs. positive) × payoff as predictors).

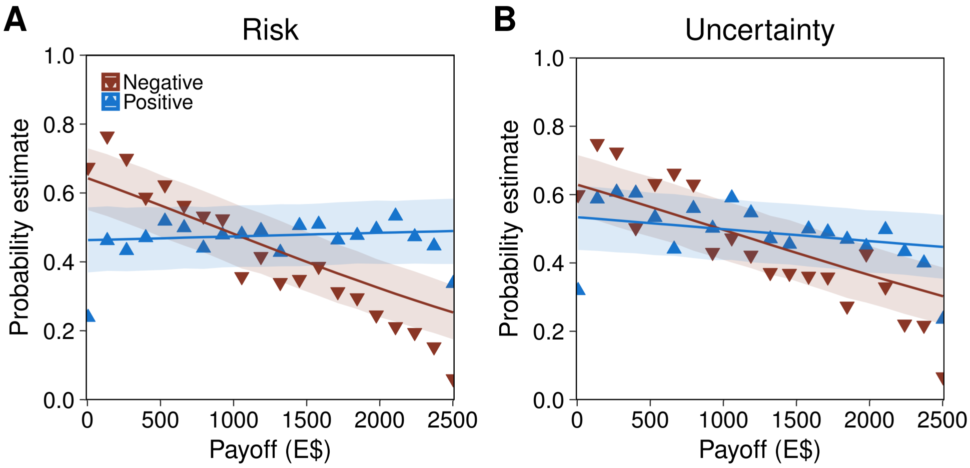


Figure 10. Payoff–probability estimation task. Participants were asked to estimate the likelihood of winning different payoffs from learning phase bets. (A) Risk: The learning phase included information about historical frequencies. (B) Uncertain: The learning phase did not include information about historical frequencies. Triangles indicate mean estimates at each payoff level. Lines (ribbons) indicate the mean (95% HDIs) of the posterior predictions.

Did individual differences in learned risk–reward structures help to predict choices of the uncertain alternative in the decision under uncertainty task? As an index of the learned risk–reward structures, we again estimated a risk–reward slope for each participant from the risk–reward task (payoff-dependent estimates with participant as a grouping factor). There was a weak but not credible association between learned risk–reward relationships and the tendency to choose the higher payoff gamble ($b_{\text{positive}} = .64$, $CI = [-.04, 1.33]$, slope × payoff × condition interaction in a fixed effects model using the negative condition as baseline, controlling for historical frequencies).

Summary

Experiment 3 revealed that participants could also learn different risk–reward relationships when the probabilities were expressed in the form of epistemic events. The evidence for learning was more pronounced when the relationship was negative than when it was positive, suggesting that the negative association may have been more in line with participants' initial 'priors'. The learned risk–reward relationship impacted subjective probability estimates about the likelihood of the event occurring. Moreover, preferences in subsequent decisions under uncertainty were to some extent environment-dependent. When participants had explicit probability information available in the learning phase—that is, when they learned under risk, choices were impacted in the low-payoff condition as if participants used both their subjective knowledge about the epistemic events and their knowledge of the risk–reward relationship to estimate subjective probabilities. We also found that there were limits to the degree to which participants used the risk–reward relationship: Subsequent choices were not impacted when participants learned under uncertainty.

General Discussion

Ecological structures between risks and rewards that are present in many real-world environments afford decision makers a solution to the problem of unknown probabilities in decisions uncertainty: Decision makers can exploit risk–reward structures to infer probabilities from the magnitude of the payoff itself (Pleskac and Hertwig, 2014). Here, we investigate two requirements of such a solution (1) that people are able to extract the environmental structure and (2) that they use the structure adaptively, as the ecological regularities can and do vary across environments. Our findings from three experiments demonstrate that people can learn risk–reward relationships from the options they experience during preferential choice. Moreover, they learned the relationships without being asked to attend to the structures (incidental unsupervised learning). Finally, the learned risk–reward relationships can guide the direction of estimates and ultimately impact preferences in decisions made under uncertainty. Next, we discuss our findings in detail and consider their broader implications for adaptive approaches to cognition.

Learning Risk–Reward Structures

Adaptive approaches to cognition seek to understand cognition within the environmental context (Anderson, 1991; Gibson, 1979; Gigerenzer et al., 1999; Marr, 1982; Shepard, 1987; Simon, 1956; Stewart et al., 2006). In the words of Herbert A. Simon (1956), "…we might hope to discover, by a careful examination of some of the fundamental structural characteristics of the environment, some further clues as to the nature of the approximating mechanisms used in decision making" (p. 130). Taking this perspective means it is equally important to identify the ecological structures to which a mind may adapt as it is to establish how the mind comes to terms with those ecological structures (Brunswik and Kamiya, 1953; Simon, 1956).

People can only *exploit* a risk–reward structure if it has entered the mind. There is good evidence that people are automatic processors of frequency information (a proxy for probabilities) (Hasher and Zacks, 1979; Zacks, 2002), and distributions of payoffs (Brown et al., 2008; Stewart et al., 2006; Olivola and Sagara, 2009; Ungemach et al., 2011). The risk–reward relationship is, however, different in that it is the joint distribution of these dimensions across different gambles. Moreover, the central goal in most decision environments is to select the best option(s), and not evaluate them based on their risk–reward relationships. People are neither explicitly informed of those relationships or learn about them from explicit feedback. Instead, it would seem that, if at all, the risk–reward relationship enters the mind via incidental, unsupervised learning(e.g., Brooks, 1978; Love, 2002; Nelson, 1984; Ward and Scott, 1987; Wattenmaker, 1991). Across three experiments, we showed that participants learn risk–reward relationships from evaluating gambles. Crucially, participants even learned there was no association between risks and rewards, especially when there was no 'local' risk–reward relationship (Experiment 2).

In addition, our results suggest that participants abstracted the relationship as a rule. The strongest support for this conclusion comes from the choice patterns and probability estimates in decisions under uncertainty. Gambles in these tasks did not perfectly map onto learning phase exemplars, yet participants' choices and probability estimates largely resembled the risk–reward rule they had learned previously. Our data cannot, however, pinpoint whether this abstraction occurs during or after encoding (Wattenmaker, 1991). Participants may have used hypotheses about what they know from risk–reward relationships to abstract a rule during encoding (Altmann et al., 1995; Wattenmaker, 1999), or represented the stimuli as exemplars and retrieved a rule from these exemplars as they needed it (Wattenmaker, 1991). Our findings point to some general factors that appear to affect how easily risk–reward structures are learned.

First, it seems that some risk–reward structures are more difficult to learn than others. Specifically, there was evidence that positive risk–reward structures are more difficult to learn than negative ones: Not all participants picked up on the positive relationship, resulting in weaker positive risk–reward estimates than in the risk–reward structure presented in the learning phase. Since people do not come across positive relationships outside the laboratory very often, they may require more evidence to acquire it. After all, in the real world, there is usually "no free lunch." They may even use prior knowledge and assume that high bonuses are unlikely in the experiment itself. Indeed, as Pleskac & Hertwig (2014) showed, people assume a negative risk–reward relationship without prior learning. Second, risk–reward structures seem to be learned more readily with some response types than with others. A comparison of Experiments 1 and 2 suggests that people are more likely to pick up the risk–reward regularity when pricing gambles one-by-one than when choosing between gambles. It is possible that pricing engages deeper processing than choosing the subjectively better option, leading to better encoding of the relationship. Another reason could be that when people choose between two nondominated gambles for all conditions, a 'local' risk–reward relationship is experienced within the choice pair (i.e., a higher payoff is associated with a lower probability relative to the other gamble). A third factor that hampers learning, as Experiment 3 illustrates, is the level of uncertainty in the choices people learn from. It could be that people need many learning trials to adapt to a new environment under uncertainty. To some extent, such conditions may reflect nonlaboratory environments (where there is both uncertainty, but also continuous learning across many, many "trials").

How Risk–Reward Structures Impact Decisions Under Uncertainty

Payoffs and probabilities are the pillars of preference. This makes decision making under uncertainty a vexing problem as one of those pillars—the probabilities—is missing. People are commonly thought to deal with this problem by intuiting subjective probabilities from their knowledge and memory (Fox and Tversky, 1998; Tversky and Fox, 1995) or by estimating statistical probabilities from samples of information (Hertwig and Erev, 2009). Our results support still another ecologically grounded solution, namely, that people estimate the missing probabilities from their immediate choice environments via their learned risk–reward relationships.

More broadly, these findings fit the general processing assumptions of a risk–reward heuristic. First, people do not seem to memorize exemplars but abstract the risk–reward relationship as a rule. Second, subsequent choice patterns speak for the subsequent use, or retrieval, of this rule, and against algebraic calculation. Taken together, these properties are consistent with the heuristic use of payoff information to estimate probabilities, rather than with the use of more complex methods. However, our experiments also identified some limitations on the use of the risk–reward relationship as a heuristic. For instance, the results of Experiment 3 demonstrate that at least under some circumstances, other information beyond the payoff information is used to infer probabilities about epistemic events. Thus, it is unclear to what extent people use the risk–reward relationship in a noncompensatory manner. This is an important aspect as ignorining other information is sometimes used as a defining characteristic of heuristics (Gigerenzer and Gaissmaier, 2011).

Regardless, the exploitation of the environmental structure has some immediate implications. One is that, as we have shown, experienced risk–reward environments can create environment-dependent preferences in decisions under uncertainty. In particular, participants in negative risk–reward environments chose the sure thing more often as payoffs increased, but the opposite occurred for participants in positive risk–reward environments. In uncorrelated environments, preferences were less extreme but still tended to track a negative risk–reward environment, perhaps reflecting the pervasiveness of negative risk–reward environments outside the lab.

This ecological dependency of preferences brings a new perspective to the proposition that preferences are constructed rather than revealed (Ariely and Norton, 2008; Lichtenstein and Slovic, 2006; Payne et al., 1992; Slovic, 1995). The construction of preferences has typically been understood as the result of people selecting a specific procedure from a larger repertoire of possible strategies to formulate a response (Brandstätter et al., 2006; Pachur et al., 2013; Payne et al., 1993; Tversky et al., 1988), the dynamic nature of information accumulation that adjusts preferences over time (Busemeyer and Townsend, 1993), or the ecological (marginal) distribution of monetary payoffs and probabilities (Birnbaum, 1992; Stewart et al., 2006, 2015; Walasek and Stewart, 2015). Here, we have shown how experiencing different risk–reward environments can result in substantial, environment-dependent preference shifts in decisions under uncertainty.

These environment-dependent preferences are not indicative of a fallacy, but an ecologically rational bet on the structure of the environment. Such a bet is more accurate than ignoring probability information altogether—for example, by using the principle of indifference and assigning equal probabilities to all outcomes (Keynes, 1921). Moreover, our results speak against overtly optimistic estimates that increase as the payoff increases, as implied by the desirability bias (Bar-Hillel and Budescu, 1995; Edwards, 1954; Krizan and Windschitl, 2007; Sharot, 2011) or the affect heuristic (Slovic and Peters, 2006). If anything, participants adapted too little to positive risk–reward environments, perhaps due to the strength and pervasiveness of negative risk–reward environments.

Conclusion

People often have to make decisions under uncertainty, when probability information is not explicitly stated. In many natural environments, risks and rewards are systematically correlated. This regularity allows people to infer the probability of a payoff from its magnitude, consistent with the use of a risk–reward heuristic. By adjusting their preferences to the respective risk–reward structure, people often manage to make highly adaptive choices under uncertainty.

References

Altmann, G., Z. Dienes, and A. Goode (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition 21*(4), 899–912.

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review 98*(3), 409–429.

Anderson, J. R. and L. J. Schooler (1991). Reflections of the environment in memory. *Psychological Science 2*(6), 396–408.

Anderson, R. B., M. E. Doherty, N. D. Berg, and J. C. Friedrich (2005). Sample Size and the Detection of Correlation–A Signal Detection Account: Comment on Kareev (2000) and Juslin and Olsson (2005). *Psychological Review 112*(1), 268–279.

Ariely, D. and M. I. Norton (2008). How actions create - not just reveal - preferences. *Trends in Cognitive Sciences 12*(1), 13–16.

Bar-Hillel, M. and D. Budescu (1995). The elusive wishful thinking effect. *Thinking & Reasoning 1*(1), 71–103.

Becker, G. M., M. H. Degroot, and J. Marschak (1964). Measuring utility by a single-response sequential method. *Behavioral Science 9*(3), 226–232.

Birnbaum, M. H. (1992). Violations of monotonicity and contextual effects in choice-based certainty equivalents. *Psychological Science 3*(5), 310–314.

Brandstätter, E., G. Gigerenzer, and R. Hertwig (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review 113*(2), 409–432.

Brooks, D. N. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch and B. B. Lloyd (Eds.), *Cognition and Categorization*, pp. 169–215. Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, G. D., J. Gardner, A. J. Oswald, and J. Qian (2008). Does wage rank affect employees' well-being? *Industrial Relations: A Journal of Economy and Society 47*(3), 355–389.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review 62*(3), 193–217.

Brunswik, E. and J. Kamiya (1953). Ecological cue-validity of 'proximity' and of other gestalt factors. *The American Journal of Psychology 66*(1), 20–32.

Busemeyer, J. R. and J. T. Townsend (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review 100*(3), 432–459.

Cleeremans, A., A. Destrebecqz, and M. Boyer (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences 2*(10), 406–416.

Cooksey, R. W. (1996). *Judgment analysis: Theory, methods, and applications*. San Diego, CA: Academic Press.

Dulany, D. E., R. A. Carlson, and G. I. Dewey (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General 113*(4), 541–555.

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin 51*(4), 380–417.

Edwards, W. (1962). Utility, subjective probability, their interaction, and variance preferences. *Conflict Resolution 6*(1), 42–51.

Erickson, M. A. and J. K. Kruschke (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General 127*(2), 107–140.

Fox, C. R. and R. T. Clemen (2005). Subjective probability assessment in decision analysis: Partition dependence and bias toward the ignorance prior. *Management Science 51*(9), 1417–1432.

Fox, C. R. and Y. Rottenstreich (2003). Partition priming in judgment under uncertainty. *Psychological Science 14*(3), 195–200.

Fox, C. R. and A. Tversky (1998). A belief-based account of decision under uncertainty. *Management Science 44*(7), 879–895.

Frensch, P. A. and D. Rünger (2003). Implicit learning. *Current Directions in Psychological Science 12*(1), 13–18.

Gelman, A. and D. B. Rubin (1992). Inference from iterative simulation using multiple sequences. *Statistical Science 7*(4), 457–511.

Gibson, J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.

Gigerenzer, G. and W. Gaissmaier (2011). Heuristic decision making. *Annual Review of Psychology 62*, 451–482.

Gigerenzer, G., R. Hertwig, and T. Pachur (2011). *Heuristics: The Foundations of Adaptive Behavior*. New York, NY: Oxford University Press.

Gigerenzer, G., U. Hoffrage, and H. Kleinbolting (1991). Probabilistic Mental Models: A Brunswikian Theory of Confidence. *Psychological Review 98*(4), 506–528.

Gigerenzer, G., P. Todd, and the ABC Research Group (1999). *Simple Heuristics That Make Us Smart*. New York, NY: Oxford University Press.

Hasher, L. and R. T. Zacks (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General 108*(3), 356–388.

Hasher, L., R. T. Zacks, K. C. Rose, and H. Sanft (1987). Truly incidental encoding of frequency information. *The American Journal of Psychology 100*(1), 69–91.

Hertwig, R. and I. Erev (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences 13*(12), 517–523.

Irwin, F. W. (1953). Stated expectations as functions of probability and desirability of outcomes. *Journal of Personality 21*(3), 329–335.

Kareev, Y. (2000). Seven (indeed, plus or minus two) and the detection of correlations. *Psychological Review 107*(2), 397–402.

Keynes, J. M. (1921). *A Treatise on Probability*. London: MacMillan and Co., Ltd.

Knight, F. H. (1921). *Risk, Uncertainty, and Profit*. New York, NY: Sentry Press.

Krizan, Z. and P. D. Windschitl (2007). The influence of outcome desirability on optimism. *Psychological Bulletin 133*(1), 95–121.

Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A tutorial with R, JAGS, and Stan, 2nd Ed.* New York,

NY: Academic Press/Elsevier.
Langer, E. J. (1975). The illusion of control. *Journal of personality and social psychology 32*(2), 311–328.
Leuker, C., T. Pachur, R. Hertwig, and T. J. Pleskac (2018). Too good to be true? Psychological responses to surprising options in risk–reward environments (Manuscript submitted for publication).
Leuker, C., T. J. Pleskac, T. Pachur, and R. Hertwig (2017). How the mind exploits risk–reward structures in decisions under risk. In G. Gunzelmann, A. Howes, T. Tenbrink, and E. Davelaar (Eds.), *Proceedings of the 39th Annual Conference Cognitive Science Society*, London, UK, pp. 2543–2548.
Lichtenstein, S. and P. Slovic (2006). *The Construction of Preferences.* New York, NY: Cambridge University Press.
Love, B. C. (2002). Comparing supervised and unsupervised category learning. *Psychonomic Bulletin & Review 9*(4), 829–835.
Luce, R. D. and H. Raiffa (1957). *Games and Decisions.* New York, NY: Dover Publications, Inc.
Marr, D. (1982). Vision: A computational approach.
McFadden, D., M. J. Machina, and J. Baron (1999). Rationality for economists? In B. Fischhoff and C. F. Manski (Eds.), *Elicitation of preferences*, pp. 73–110. Dordrecht: Springer.
Nelson, D. G. K. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning and Verbal Behavior 23*(6), 734–759.
Newsome, J. (2016, March 7). Man wins $291 million in lottery, brother wins $7. *CNN*.
Olivola, C. Y. and D. M. Oppenheimer (2008). Randomness in retrospect: Exploring the interactions between memory and randomness cognition. *Psychonomic Bulletin and Review 15*(5), 991–996.
Olivola, C. Y. and N. Sagara (2009). Distributions of observed death tolls govern sensitivity to human fatalities. *Proceedings of the National Academy of Sciences 106*(52), 22151–22156.
Pachur, T., G. G. Hertwig, Ralph, and E. Brandstätter (2013). Testing process predictions of models of risky choice: a quantitative model comparison approach. *Frontiers in Psychology 4*, 646.
Pachur, T., R. Hertwig, and F. Steinmann (2012). How do people judge risks: Availability heuristic, affect heuristic, or both? *Journal of Experimental Psychology: Applied 18*(3), 314–330.
Payne, J. W., J. R. Bettman, and E. J. Johnson (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology 43*(1), 87–131.
Payne, J. W., J. R. Bettman, and E. J. Johnson (1993). *The Adaptive Decision Maker.* New York, NY: Cambridge University Press.
Peirce, J. W. (2007). Psychopy - Psychophysics software in Python. *Journal of Neuroscience Methods 162*(1-2), 8–13.
Perkovic, S. and J. L. Orquin (2017). Implicit Statistical Learning in Real-World Environments Leads to Ecologically Rational Decision Making. *Psychological Science 29*(1), 34–44.
Pleskac, T. J. and R. Hertwig (2014). Ecologically rational choice and the structure of the environment. *Journal of Experimental Psychology: General 143*(5), 2000–2019.
Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior 6*(6), 855–863.
Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General 118*(3), 219–235.
Seger, C. A. (1994). Implicit learning. *Psychological Bulletin 115*(2), 163–196.
Shanks, D. R. and M. F. St. John (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences 17*(3), 367–395.
Sharot, T. (2011). The optimism bias. *Current Biology 21*(23).
Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior 6*(1), 156–163.
Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science 237*(4820), 1317–23.
Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review 63*(2), 129–138.
Slovic, P. (1995). The construction of preference. *American Psychologist 50*(5), 364–371.
Slovic, P., M. L. Finucane, E. Peters, and D. G. MacGregor (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis 24*(2), 311–322.
Slovic, P. and E. Peters (2006). Risk perception and affect. *Current Directions in Psychological Science 15*(6), 322–325.
Stan Development Team (2016). RStanArm: Bayesian applied regression modeling via Stan (version 2.9.0-4). R package version 2.13.1.
Stewart, N., N. Chater, and G. D. A. Brown (2006). Decision by sampling. *Cognitive Psychology 53*(1), 1–26.
Stewart, N., S. Reimers, and A. J. Harris (2015). On the origin of utility, weighting, and discounting functions: How they get their shapes and how to change their shapes. *Management Science 61*(3), 687–705.
Todd, P. M. and G. Gigerenzer (2007). Environments that make us smart: Ecological rationality. *Current Directions in Psychological Science 16*(3), 167–171.
Tversky, A. and C. R. Fox (1995). Weighing risk and uncertainty. *Psychological Review 102*(2), 269–283.
Tversky, A. and D. Kahneman (1986). Rational choice and the framing of decisions. *Journal of Business 59*(4), S251–S278.

Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty 5*(4), 297–323.

Tversky, A., S. Sattath, and P. Slovic (1988). Contingent weighting in judgment and choice. *Psychological Review 95*(3), 371–384.

Tversky, A. and P. Wakker (1995). Risk attitudes and decision weights. *Econometrica 63*(6), 1255–1280.

Ungemach, C., N. Stewart, and S. Reimers (2011). How incidental values from the environment affect decisions about money, risk, and delay. *Psychological Science 22*(2), 253–260.

Wakker, P. (2010). *Prospect Theory for Risk and Ambiguity.* Cambridge, UK: Cambridge University Press.

Walasek, L. and N. Stewart (2015). How to make loss aversion disappear and reverse: Tests of the decision by sampling origin of loss aversion. *Journal of Experimental Psychology: General 144*(1), 7–11.

Ward, T. B. and J. Scott (1987). Analytic and holistic modes of learning family-resemblance concepts. *Memory & Cognition 15*(1), 42–54.

Wattenmaker, W. D. (1991). Learning modes, feature correlations, and memory-based categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition 17*(5), 908–923.

Wattenmaker, W. D. (1999). The influence of prior knowledge in intentional versus incidental concept learning. *Memory & Cognition 27*(4), 685–698.

Whittlesea, B. W. (1987). Preservation of specific experiences in the representation of general knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition 13*(1), 3–17.

Windschitl, P. D., A. R. Smith, J. P. Rose, and Z. Krizan (2010). The desirability bias in predictions: Going optimistic without leaving realism. *Organizational Behavior and Human Decision Processes 111*(1), 33–47.

Zacks, R. T. (2002). Frequency processing: A twenty-five year perspective. In P. Sedlmeier and T. Betsch (Eds.), *Frequency processing and cognition*, pp. 21–36. New York, NY: Oxford University Press.

3 | Too good to be true? Psychological responses to surprising options in risk–reward environments

Leuker, C., Pachur, T., Hertwig, R. & Pleskac, T.J. *Under Review.*

Abstract

In decisions under risk, options that sound too good to be true often *are* too good to be true. But how do decision makers detect if something is too good to be true, and are what their psychological responses to such options? We argue that they compare options against a regularity present in many natural domains, namely an inverse relationship between risks and rewards (Pleskac and Hertwig, 2014), according to which highly attractive options are very unusual. We investigated behavioral and physiological responses to deviations from such regularities as participants evaluated risky options. In two experiments ($N = 183$), participants first priced monetary gambles drawn from environments in which risks and rewards were negatively correlated, positively correlated, or uncorrelated. In later trials, they were presented with "surprising" gambles that deviated from the respective environment's risk–reward structure. Pricing, response times and (in Experiment 2) pupil dilation were recorded. In both experiments, participants took more time when responding to "surprising" than to "expected" options. This result was most pronounced for "surprisingly good" options, that offered higher expected values compared to the other gambles in the set. Moreover, these surprisingly good options were associated with an increase in pupil size. These results suggest that people's evaluations of risky options are based not only on their payoffs and probabilities, but also on the extent to which they fit the risk–reward structure of the environment.

Acknowledgments: We thank Chantal Wysocki and Lisa-Marie Jirak for assistance with data collection.

Introduction

On December 25th, 2017, Nicole Coggins and her mother-in-law bought 100 lottery tickets for the South Carolina Holiday Cash Add-a-Play lottery for $1 each. Astonishingly, every ticket was a winner, giving them a total of $18,000 in winnings. Wade Crenshaw, who was working behind the cash register at a convenience store that day, noticed that more and more people were asking to buy Add-A-Play tickets. He later said, "It was weird, everybody [was] winning so much. I didn't know if they were doing some kind of Christmas special" (Fortin, 2017). If you also think that this sounds "too good to be true," you are right: Except for a select few individuals who were able to collect a total windfall of about $1.7 million when most of the lucky winners went to cash in their prizes, the machine deemed their tickets to be invalid (Fortin, 2018; Harrington, 2017). As it later emerged, the many winning lottery tickets were not a Christmas special; they were due to a computer glitch.

As the reaction of the cashier illustrates, people are usually aware that higher rewards are more unlikely than smaller rewards, a regularity present across many monetary and nonmonetary domains in the environment (Pleskac and Hertwig, 2014). This link between risks and rewards (or probabilities and payoffs) affords people an ecological structure that they may use when making decisions (Brunswik, 1943). One such use is that during decisions under uncertainty, when people know the possible payoffs of an option but not their probabilities. When probabilities are unknown, people often exploit the risk–reward relationship to infer the probabilities from the magnitude of the payoffs (Leuker et al., 2018; Pleskac and Hertwig, 2014). They may also use the relationship to infer payoffs from probabilities (Skylark and Prabhu-Naik, 2018). Here we examine if and how an environmental relationship between risks and rewards results in options being perceived as surprising. Thus, our focus here is on *decisions under risk*—that is, decisions in which both payoffs and probabilities are explicitly stated.

Responses to surprising stimuli have been studied in other areas of psychology and neuroscience. Researchers using what has become known as the oddball paradigm (Herrmann and Knight, 2001; Huettel and McCarthy, 2004; Picton, 1992; Squires et al., 1975) have examined whether participants automatically detect surprising perceptual stimuli in a sequence (called "deviants" or "oddballs"), without being explicitly instructed to process a particular sequence or deviations from it. For example, in a typical auditory oddball paradigm, participants are presented with a sequence of standards (s) and a small subset of deviants (d) that can differ in volume, duration, or pitch (e.g., s–s–s–s–s–s–d–s–s–d–s–s–s). Typically, the focus in these studies is an electrophysiological brain response to the deviants called mismatch negativity. Prior research showed that deviants are detected even when participants' attention is not focused on the stimuli (for reviews, see Näätänen, 2007; Garrido et al., 2009).

Risky choice differs from these rather simple perceptual tasks. One difference is that during risky choice people typically have to attend to and process multiple attributes (i.e., risks and rewards). This impacts what constitutes standards and deviants. In risk–reward environments, there are many different combinations of probabilities and payoffs that are consistent with a particular risk–reward structure. For instance, $10 with $p = .9$; $20 with $p = .8$; and $30 with $p = .7$ are all in line with a negative risk–reward structure in which risks and rewards are inversely related. The multiple attributes also mean that there can be different relationships between attributes and thus one option can be a deviant or surprising in one environment (e.g., when there is a negative relationship between risks and rewards), but a standard in another environment (e.g., when there is a positive relationship). A final difference is that risky choice is based on processing value-based information. Since options in risky choice have an intrinsic value, deviants from a risk–reward structure can either offer a "surprisingly good" combination of risks and rewards or a "surprisingly bad" one.[1] In the perceptual domain, by contrast, the direction in which a stimulus deviates from the norm should usually matter less, or not at all.

Physiological responses to surprising stimuli have been investigated in value-based paradigms from a different perspective. Specifically, it has been shown that both human and animals are sensitive to "risk-prediction errors"—that is, to a mismatch between expected and experienced probabilitistic outcomes (Bossaerts, 2010; O'Neill and Schultz, 2013; Preuschoff et al., 2008). One key finding is that mispredictions in either direction (probability is low but a reward is still obtained, or probability is high but no reward is obtained; see Preuschoff et al., 2011) seem to be associated with increased pupil dilation.[2] In deviations from risk-reward structures, the nature of surprise is somewhat different, as it is based not on the outcome, but on the mismatch between the properties of the current option (probabilities and payoffs) and its environment. In addition, detecting surprising options in risk–reward environments does not necessitate feedback about whether or not an outcome is obtained. We return to these conceptual (and possibly functional) differences between our experiments and previous studies of surprise in the discussion section. Next we outline three main hypotheses regarding how people respond to and process surprising options in risk–reward environments.

Overview of Experiments and Hypotheses

In two experiments, we investigated behavioral (Experiments 1 & 2) and physiological (Experiment 2) responses to options that deviated from a previously learned risk–reward environment and were thus surprising. Briefly, between participants, the structure of the environment was manipulated to be negatively correlated, positively correlated, or uncorrelated (Figure 1). In both experiments, participants indicated the price at which they would be willing to sell a monetary gamble of the form "p chance of winning x, otherwise nothing."

There were two classes of gambles. *Environment gambles*, denoted as black circles in Figure 1, defined the structure of the environment in a given condition. *Test gambles*, denoted as triangles in Figure 1, were common to all conditions and used to test our hypotheses. Depending on the environment they were interspersed in, the test gambles belonged to one of three different groups. One subset of the test gambles were *surprising gambles* (red triangles). These gambles were inconsistent with the risk–reward structure when it was present. A second subset of gambles were *expected gambles* (blue triangles). These gambles were consistent with the risk–reward structure when it was present (in negative or positive risk–

[1] In this article we are concerned with peoples' evaluations of options in the gain domain; thus, "surprisingly bad" options still involve a gain (but a small and unlikely one).

[2] Reward-prediction errors, in contrast, involve a mismatch between expected and experienced rewards. In contrast to risk-prediction errors, the neurobiological correlates differ depending on the *direction* of the misprediction: Rewards exceeding expectations lead to a positive reward-prediction error and an increase in dopaminergic firing; rewards worse than expectations lead to a negative reward-prediction error and smaller dopaminergic firing rates (Schultz, 2002; Schultz et al., 1997).

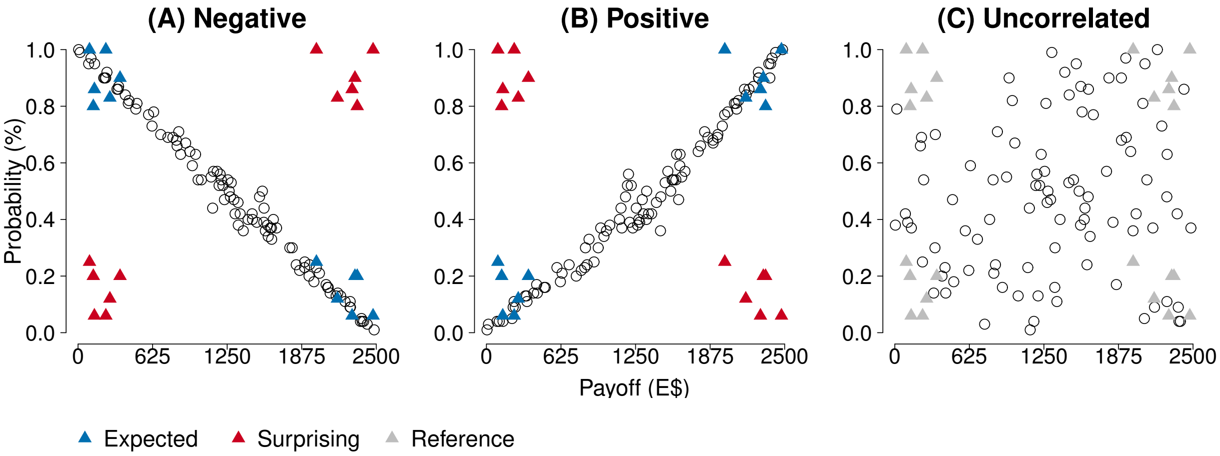


Figure 1. Prototypical gambles used in the two experiments. Environment gambles (black circles) were drawn from one of the three risk–reward environments. A set of test gambles, which was common to all three conditions, was randomly interspersed after two-thirds (Experiment 1) or one-third (Experiment 2) of the trials. Test gambles are color-coded by gamble type: Test gambles shown in blue were consistent with the risk–reward relationships in a condition and could therefore be expected (in the following referred to as "expected gambles"). Test gambles shown in red were inconsistent with risk–reward relationships in a condition and were therefore surprising (in the following referred to as "surprising gambles"). Test gambles shown in light gray (panel C) served as reference gambles. Here, participants were unlikely to have any expectations about particular risk–reward relationships, due to the environment gambles being uncorrelated (in the following referred to as "reference gambles").

reward environments). Finally, test gambles that were neither expected nor surprising (in the uncorrelated condition) are referred to as *reference gambles* (grey triangles). Participants were only exposed to test gambles after sufficient exposure to the environment gambles (see procedures for each experiment for details). Our hypotheses refer to a comparison between either the surprising and expected gambles or the surprising and reference gambles.[3]

Participants were not instructed to pay attention to the underlying risk–reward structure in either experiment. We examined how pricing and response times (RTs) varied as a function of whether the gambles were surprising or expected, relative to the reference gambles in the uncorrelated condition. In Experiment 2, in addition to the behavioral responses, we tracked pupil size in response to surprising, expected, and reference gambles.

Hypotheses

Pricing

Our first set of hypotheses focuses on the prices people provide in response to surprising options. Prices may deviate from gamble's expected values in the direction of what can be expected: When it was learned that high payoffs usually co-occur with low probabilities, prices given for high payoff/high probability options may undershoot the gambles' expected values. Conversely, when it was learned that a high payoffs usually co-occur with high probabilities, prices given to high payoff/low probability options may overshoot the gambles' expected values. Alternatively, surprising options could just lead to more error in entering

[3]Technically, there was another subset of test gambles that could be called "average" gambles. They gambles appeared in each environment and were in the mid-range of the payoffs and probabilities. Because they fit all risk–reward environments equally well but had exactly the same characteristics, these gambles were used as control stimuli to examine condition-dependent differences. Briefly, as expected, there were no condition–dependent differences for these gambles.

the price of an option as these options have not been encountered in the past. In this case, there should be no systematicity in the direction in which the prices for surprising options deviate from their expected values.

Response times

Our second set of hypotheses deals with the processing time people allocate to surprising options. One hypothesis was that increasing familiarity with a risk–reward structure of an environment should accelerate the processing of subsequent options consistent with that structure, and decelerate the processing of subsequent options inconsistent with that structure. Such a pattern would be consistent with findings in the domain of event sequence learning (i.e., responses to stimuli in different locations that follow a sequence, e.g., 4–3–2–1–4–3–2–1). Here, longer RTs for stimuli inconsistent with a learned sequence are taken as direct evidence that a sequential stimulus structure has been learned (Rüsseler and Rösler, 2000).

Moreover, and not mutually exclusive to the first hypothesis, response times may vary as a function of the *direction* of surprise. As noted earlier, a unique feature of surprising value-based stimuli is that options can be surprisingly good or surprisingly bad. Prior research found that people may have a mechanism in place that "prevent[s] impulsive responding due to the presence of high value options" (Cavanagh et al., 2014, p. 2) (also see Frank et al., 2007). By extension, participants may respond more slowly to surprisingly good options than to surprisingly bad ones. Note such a mechanism would also lead to response times increasing as a function of absolute value, rather than as a function of surprise.

Lastly, someone who has learned that risks and rewards are almost perfectly correlated may exploit this statistical regularity (Simon, 1956) and *infer* one attribute from the other — payoffs from probabilities or probabilities from payoffs — instead of looking up both attributes. This would prevent a decision–maker from detecting a surprising stimulus altogether. In this case, response times would not differ as a function of surprise but there should be strong deviations in the pricing of the options towards what is being expected.

Pupil dilation

Inspired by paradigms investigating feedback-based "risk-prediction errors" (Preuschoff et al., 2011), in Experiment 2 we modified the pricing paradigm used in Experiment 1 to measure pupil dilation as participants inspected the properties of a gamble. We did so by sequentially presenting the payoff first, and then the probability for each gamble. We predicted that participants would be surprised by options for which the probability information deviated from the learned risk–reward environment, and that surprise would only manifest itself as the probability information was revealed. That is, similar to Preuschoff et al. (2011), we hypothesized that participants would show greater pupil dilation when an option turned out to be surprising (i.e., to have an unexpected probability, given the payoff), but not in response to just seeing high or low payoffs.

Pupil dilation may be linked to surprising options for different reasons. Recent research has shown that pupil dilation is associated with an increase in the amount of information that is collected when people are presented with very similar options (see Cavanagh et al., 2014, for choices and pupil dilation modeled as a drift diffusion process). This increase in the amount of information collected results in a more rigorous evaluation of the alternatives and longer RTs. A similar mechanism might emerge for surprising options in risk–reward environments, in which a more rigorous evaluation of surprising options may be linked to both longer RTs and increases in pupil size. Thus, participants may scrutinize more carefully only those surprising options that offer a surprisingly high expected value and not those that offer a surprisingly low one (e.g., due to less impulsive responding in the presence of high-value options; Frank et al., 2007).

Experiment 1: Behavioral Responses To Surprising Risk–Reward Combinations

In Experiment 1, participants priced monetary gambles drawn from negative, positive, or uncorrelated risk–reward environments. Our main question was how prices and RTs would differ when participants encountered gambles that represented surprising risk–reward combinations relative to expected risk–reward combinations (i.e., in the positive and negative environments) and relative to reference gambles (i.e., in the uncorrelated environment, when gambles could not be compared against any risk–reward regularity). In the experiment, we also collected data on how these same participants made judgments and decisions about uncertain events and how exposure to different risk–reward environments impacted these responses. This data is reported in Experiment 2 of Leuker et al. (2018).

Method

Participants

We recruited 90 participants (53 females, age 24.7 years, $SD = 4.1$ years, proportion students = .72) from the participant pool maintained at the Max Planck Institute for Human Development. Each participant completed the experiment (duration 65 min) in exchange for a show-up fee of €10 and a performance-based bonus (€1.99 – €7.82).

Stimuli

Participants priced monetary gambles of the form “p chance of winning x, otherwise nothing.” All payoffs were expressed using an experimental currency, $E\$$, with a disclosed conversion rate of $2500E\$ = €1$). We used an experimental currency to minimize the impact of outside norms associated with specific currencies on the experiments. For the negative condition, these gambles were constructed as follows: 150 random payoffs were drawn from a uniform distribution with the range $1.01 - 2500$. The probabilities for each payoff were set such that they were inversely related to the payoff x ($p = 1 - x/2500$). We jittered payoffs and probabilities by adding normally distributed noise with a standard deviation of 0.1 to both the logit transformation of the probabilities and the logit transformation of normalized payoffs. We then transformed those perturbed values back to the scales used in the experiment. For the positive condition, we used the same gambles as in the negative condition but reversed the order of the probabilities. For the uncorrelated condition, we randomly linked probabilities and payoffs. This approach controlled for the

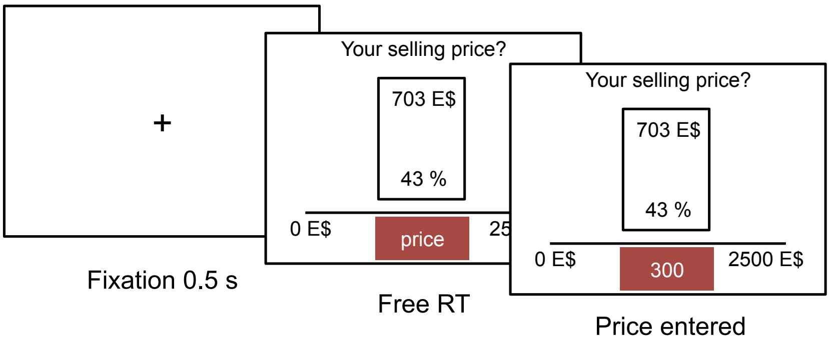


Figure 2. Task. In Experiment 1, participants saw a gamble and priced it in their own time. The price was set by moving an arrow along a rating scale and confirming the price by clicking on it.

marginal distribution of payoffs and probabilities across conditions (as marginal distributions of payoffs and probabilities may influence choice; see Birnbaum, 1992; Stewart et al., 2006).

In addition to these 150 condition-dependent gambles, participants also priced 22 gambles that were common to each of the three conditions, yielding 172 gamble stimuli per condition. Specifically, we included 12 test gambles at the extreme ends of the payoff–probability distribution space. These consisted of 3 high payoff/high probability, 3 high payoff/low probability, 3 low payoff/high probability, and 3 low payoff/low probability gambles; see triangles in Figure 1). Payoffs were random draws between $1E\$-500E\$$ (low) and $2000E\$-2500E\$$ (high). Probabilities were random draws between $0.01-.2$ (low) and $.8-1.0$ (high). Payoffs and probabilities were factorially combined to obtain the four payoff/probability combinations. The test gambles were interspersed after 100 environment-only trials (i.e., in the last third of the trials). We also included 10 mid-range payoff/mid-range probability gambles (payoffs around $1250E\$$, probabilities around .5). These gambles were equally consistent with all conditions because they were not linked to either high or low (i.e., extreme) payoffs (not depicted in Figure 1). We added them to be able to study condition-dependent differences beyond gambles being surprising, expected, or reference gambles and controlling for expected value differences. Briefly, average payoff/average probability gambles were evaluated similarly across all three risk–reward conditions.

Procedure

Participants indicated their willingness to sell (WTS) for one gamble at a time (see Figure 2), taking self-paced breaks between five blocks. The task was presented in the form of a game show called "Keep or sell?" ("Behalten oder Verkaufen?"). To motivate participants to indicate their true valuations of the gambles, we implemented a Becker-DeGroot-Marschak auction (Becker et al., 1964) as follows: Participants entered a price at which they would be willing to sell each gamble by moving the mouse along a rating scale ($0E\$ - 2500E\$$) and confirming the value with a click. To incentivize the task, the experimenter informed participants that 10 gambles would be randomly selected at the end of the experiment. For those 10 gambles, the experimenter then offered a randomly generated buying price between 0 and the absolute payoff in that gamble. If the experimenter's buying price exceeded the participant's selling price, participants sold the gamble and earned the buying price. If the participant's selling price exceeded the experimenter's buying price, the gamble was played out (e.g., 50% chance of $380E\$$). The dominant strategy in this task is to price a gamble based on its subjective value: Setting higher prices can prevent participants from selling unattractive gambles; setting lower prices can lead to them selling attractive gambles under value. In other words, the prices should approximate participants' certainty equivalents for the gambles. Experiments were coded in PsychoPy (Peirce, 2007).

Statistical Analyses

We used Bayesian estimation techniques (Kruschke, 2015). Specifically, we applied Bayesian generalized linear mixed models using Stan in R for regression analyses with the rstanarm package (Stan Development Team, 2016). All regression models used trial-level data and participant as a grouping factor. We ran three chains using Markov Chain Monte Carlo sampling to draw from posterior distributions of parameters. Depending on model complexity, we ran $10,000 - 30,000$ samples per chain (to ensure an effective sample size of $>10,000$ for each coefficient) and set a burn-in of 500 samples. We investigated (convergence of) our posteriors through visual inspection and the Gelman–Rubin statistic (Gelman and Rubin, 1992). In general, we report the mean of the posterior distribution of the parameter or statistic of interest and two-sided 95% equal tail credible intervals (CI) around each value.

In all analyses, we examined how condition-dependent expectations modulated the behavioral measure of interest (deviations from prices, deviations from individual median RTs, pupil dilation). We modeled all gamble types simultaneously and used the low payoff/low probability gambles as baseline gambles. We used both "expected" and "reference" gambles as the baseline comparison to surprising gambles in two different sets of regressions. We summarize the main results in the manuscript; beta coefficients using each of these baseline conditions and beta coefficients for each of the factorial payoff/probability combinations can be found in the Supplementary Material. As the RT data were slightly right-skewed, we normalized RTs using a log-transformation of the data before running our analyses. For better interpretability, we report and plot parameters and credible intervals from regression models using untransformed data. Qualitatively, the conclusions were identical for log-transformed and untransformed data (see Open Science Framework).

Results

We excluded trials in which participants indicated prices that exceeded the payoff offered in the gamble by more than 100E\$, as this suggests lack of attention to the task. We also removed trials in which RTs deviated by $+/-3SD$ from an individual's median, assuming that these RTs are unlikely to reflect cognitive processing in a specific trial. In total, we removed 4.6% ($718/15,480$) of the trials across all participants. We analyzed our data with and without these excluded trials and obtained qualitatively very similar results. In what follows, we note when these exclusions led to qualitatively different results.

Prices

We expressed prices as deviations from a gambles' expected value ($gamble_{price} - gamble_{EV}$). Thus, deviations can range from $-2500E\$$ to $2500E\$$; with 0 indicating perfect alignment between prices and the gambles' expected values. We first compared the extent to which prices deviated from expected values across conditions. There were no credible differences between the conditions in the extent to which prices deviated from to gamble's expected values (using the environment gambles; for estimates see Supplementary Table B1).

Focusing on the test gambles, we hypothesized that participants would provide prices deviating from the gambles' expected values when risk–reward combinations were surprising. Specifically, we hypothesized a deviation in the direction of the probability expected, given the structure of the environment. However, as Figure 3 shows, there were no reliable pricing differences as a function of whether the gamble was surprising vs. expected ($b_{deviation} = 3.44E\$$, $CI = [-24.91E\$, 31.85E\$]$, surprising vs. expected across gamble types); or whether a gamble was surprising vs. a reference gamble ($b_{deviation} = -11.59E\$$, $CI = [-49.15E\$, 26.20E\$]$, surprising vs. reference across gamble types). Moreover, as a comparison between panels in Figure 3 shows, prices did not credibly vary as a function of different payoff-probability combinations across conditions.

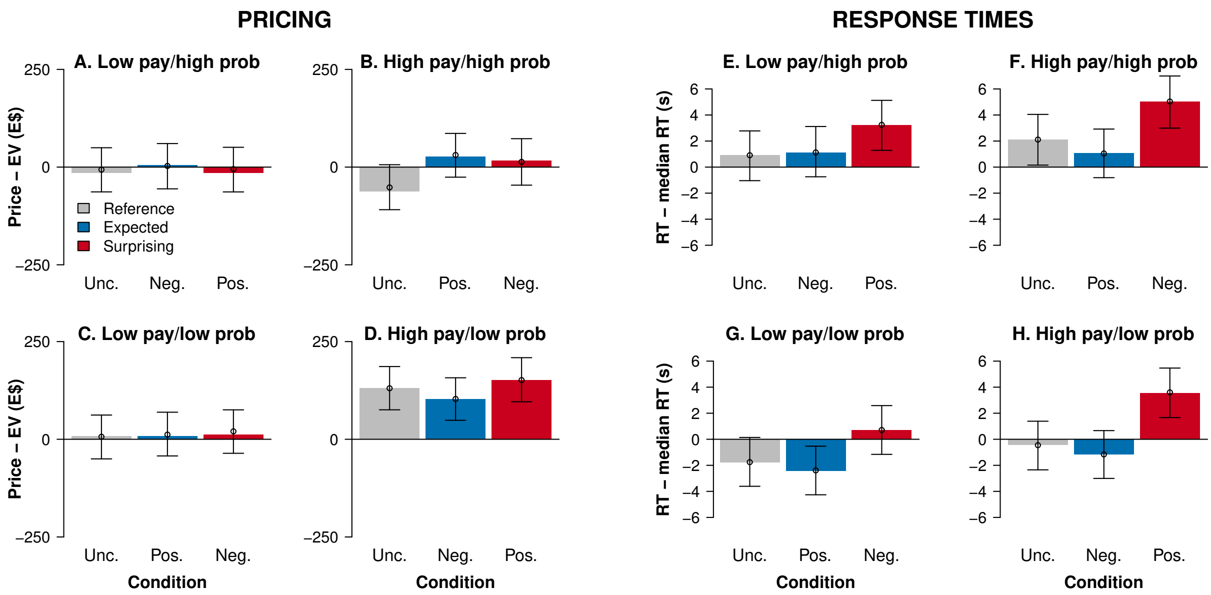


Figure 3. Behavioral results for the test gambles interspersed in Experiment 1. (A–D) Pricing. We found little to no difference in how well prices were adjusted to the expected values of the surprising gambles. (E–H) Response times. Participants slowed down when options were surprising. RTs are expressed as deviations from individual median RTs across trials ($RT_{trial} - RT_{ind.Md}$). Black dots and error bars represent the mean and the 95% credible interval of the posterior predictive distribution.

Response times

The mean RT across all gamble types was $14.17s$ per trial in the uncorrelated condition ($CI = [12.40s, 15.91s]$) and credibly lower in the negative ($M = 10.3s$, $b = -3.83s$, $CI = [-6.30s, -1.30s]$), but not in the positive conditions ($M = 13.0s$, $b = -1.17s$, $CI = [-1.37s, -0.38s]$, both models using the uncorrelated condition as a baseline).[4]

How did RTs differ when gambles were surprising? To address this question and to control for individual differences in baseline RTs, we used the difference between the observed RT and the median RT as an indicator of change in processing time ($RT_{trial} - RT_{ind.Md}$). As Panels E–H in Figure 3 show, participants

[4]When all trials were included in the analysis, participants' average RT decreased to $10.8s$ in the uncorrelated condition ($CI = [8.54s, 11.57s]$) and was again slightly faster in the correlated conditions, $b_{neg} = -1.18s$, $CI = [-4.44s, -0.008s]$, $b_{pos} = -1.94s$, $CI = [-5.15s, -0.85s]$.

slowed down in response to surprising gambles (red bars). On average, they spent around 3.5s longer on surprising gambles than on expected gambles (b = 3.47s, CI = [2.52s, 4.43s]). The RTs to the reference gambles from the uncorrelated environment were statistically indistinguishable from those to expected gambles. Consequently, the effects involving surprising gambles were qualitatively the same as when using these gambles as a baseline (see Supplementary Table B3).

When RTs were broken down by specific payoff–probability combinations, participants were 2.4s faster in responding to low payoff/low probability gambles in all conditions, irrespective of whether the gambles were surprising or not (b = −2.39s, CI = [−3.81s, −0.96s]). In comparison, all participants took around 3s longer when responding to the high payoff/high probability gambles (b = 3.43s, CI = [1.56s, 5.32s]). Responses slowed down by about another second when these high payoff/high probability gambles were also surprising (i.e., in the negative condition; see Figure 3, panel F), but this difference was not credible (b = .79s, CI = [−1.92s, 3.49s]). All results were qualitatively the same in a model using normalized, log-transformed RTs; or when using the "reference" gambles as a baseline (see Supplementary Table B3).

Summary of Experiment 1

Experiment 1 suggests that people form expectations about risk–reward structures that reflect the experienced environments. This was evident in participants' response times for options that did not match these expectations. Specifically, they slowed down in response to surprising options. At the same time, prices were equally adjusted to the gambles' expected values, irrespective of gambles being surprising or not. In addition to the overall RT effects, participants' processing of the options depended on the gambles' expected values: Differences in RTs were more pronounced for surprising high payoff/high probability gambles than for surprising low payoff/low probability gambles, to which participants responded faster in general. The high payoff/high probability gambles were surprising in the negative condition, which was characterized by otherwise similar and smaller expected values (all below 700E\$), suggesting that "EV surprise" may amplify "risk–reward surprise" by affecting participants' perceptions of possible rewards in a given experiment (for a related result see Ludvig et al., 2014, 2018). Notably, these surprise effects emerged irrespective of any feedback on whether or not an outcome was obtained.

Experiment 2: Pupillometric and Behavioral Responses to Surprising Risk–Reward Combinations

In this experiment, we sought to replicate and extend the findings of Experiment 1. We extended the findings by using pupil dilation as a physiological measure of surprise. We hypothesized that pupil dilation would increase when participants were presented with gambles offering a surprising combination of risk and reward. In order to dissociate pupillary responses to high payoffs and responses to particular payoff–probability combinations, for each gamble we first presented payoffs and added probability information after 2s.

Methods

We adapted the methodology from Experiment 1 to include an eye-tracking component that allowed us to measure pupillary responses to surprising options. We outline key differences below.

Participants

Ninety-three (55 female) participants (age $M = 25.6$ years, $SD = 3.7$ years) from the participant pool at the Max Planck Institute for Human Development, Berlin, completed the experiment (duration 75 min). All participants were paid a fixed rate of €12 plus a bonus based on their performance (€3.53–€11.67).[5]

Stimuli

We reduced the number of stimuli to 90 condition-dependent gambles and increased the number of gambles common to each of the three conditions.[6] As test gambles, we created six gambles for each payoff–probability combination. Overall, these procedures resulted in 120 gambles. The test gambles were interspersed after 40 environment-only trials. Again, we also added six average payoff/average probability gambles (not depicted). As these gambles fit equally well within each condition and were not linked to either high or low (i.e., extreme) payoffs, they were used to study condition-dependent differences beyond gambles being surprising, expected, or reference gambles (they are not depicted in Figure 1).

Procedure

Experiment 2 differed from Experiment 1 in that the payoff and probability information appeared sequentially: After a fixation cross, the payoff appeared for $2s$, followed by the probability for another $2s$. After a blank screen ($2s$), the screen automatically switched to a rating scale. Participants entered the prices they were willing to sell the gamble for by moving the mouse along this rating scale ($0E\$ - 2500E\$$) and clicking on the value to confirm (free RT). This sequential presentation of only payoffs and then payoffs and probabilities combined allowed us to analyze pupillary responses to payoffs only and pupillary responses to the joint presentation of payoffs and probabilities (Figure 4). To control for the pupillary light reflex, we matched the gambles' luminance to the background of the screen by defining stimuli colors in the Derrington, Krauskopf, and Lennie color space (see Derrington et al., 1984; MacLeod and Boynton, 1979). Gambles were presented in orange on a gray background with the same luminance (for clarity, these colors are not shown in the figure).

[5] After the pricing task reported here, participants completed a two-alternative forced choice task. The respective data and analyses are reported elsewhere (Leuker et al., 2017). The reported duration includes the additional choice tasks.

[6] We did this for the sake of time participants needed to spend (since presenting payoffs and probabilities in a fixed sequence takes more time than presenting the full gamble immediately), while simultaneously obtaining data from a larger set of test gambles.

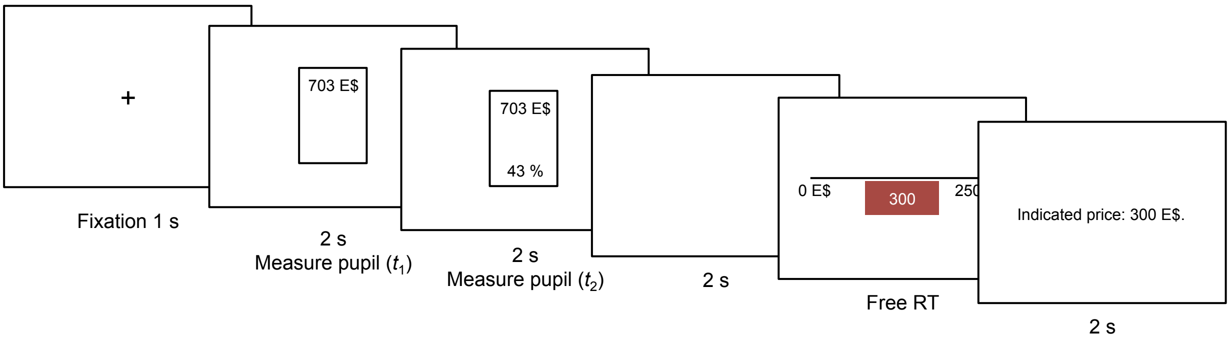


Figure 4. In Experiment 2, participants priced gambles after seeing the payoff and probability of each gamble in a fixed sequence. After 4s in total, the gamble disappeared from the screen. We added a blank screen to achieve sufficient spacing between the critical stimuli (i.e., a gamble's risk–reward relationship) and the participant's manual response. Subsequently, participants had as much time as they wanted to indicate a price for the gamble on a rating scale. We measured pupil dilation as a response to a gamble's payoff alone (t_1) and as a response to its payoff and probability combined (t_2).

Eye Tracking

We collected binocular eye tracking data with an EyeTribe tracker, sampled at 60Hz. The experiment was implemented in PsychoPy 1.83.01 (Peirce, 2007) and the eye-tracking interface PyTribe (Dalmaijer et al., 2014). Before the task, each participant's eye movements were calibrated using the EyeTribe UI with a 9-point grid ($<$ 0.7 degrees of visual angle). Participants were seated approximately 60 cm from the screen with their chin on a chinrest affixed to the table, in a room with negligible ambient light. We obtained pupil size from the left and the right eye (arbitrary units, measured at every Hz).

Eye-Tracking Analyses

Pupillary data were preprocessed as follows. We used EyeTribe's default settings to detect fixations and removed saccade data, because pupillary responses during (and even before) saccades differ systematically from those during fixations (Mathôt et al., 2016). Trials were discarded if pupil size deviated more than 3SDs from a participant's median pupil size and if it was outside plausible values (range $[10, 40]$ in arbitrary units given by Eyetribe). This procedure removed blinks (rows with values $[0, 0]$) and measurement error. We smoothed the data using a lowess filter, and we averaged the pupil size of the left and the right eye. We removed trials with fewer than 15 samples, which indicates extremely poor eye tracking (one would expect 240 samples, minus a few blinks, per 2s period). This applied to a small proportion of all trials: .07 of all trials at t_1 (when reward was shown); .05 of all trials at t_2 (when reward and probability were shown).

To facilitate comparisons across participants and consistent with the literature (Preuschoff et al., 2008; Cavanagh et al., 2014; Fiedler and Glöckner, 2012), we analyzed pupillary signals aligned to a baseline pupil size. As a baseline signal, we used the offset of stimulus presentation (median value in the first 0.1s of a trial).[7] We did so by subtracting the signal at each time point from the baseline signal and then dividing by the baseline signal, resulting in a percentage change relative to the stimulus onset.

[7]We used this baseline to obtain a similar measure of pupil changes for both t_1 and t_2. Often the fixation cross time is used as a baseline; in our paradigm, however, the fixation cross only preceded stimulus appearance at t_1; see Figure 4.

Pupillary responses to psychologically relevant stimuli are assumed to occur after approximately $1s$, and are therefore conceptually different from the pupillary light reflex that occurs after milliseconds (Gagl et al., 2011; Van Steenbergen and Band, 2013). As in previous research using pupillary responses in the context of choice (Cavanagh et al., 2014), we therefore set an a priori region of interest from 1 to $2s$ poststimulus at t_1 (payoff visible on screen) and t_2 (payoff and probability visible together on screen) for our statistical analyses (see Figure 1). We obtained the median percentage change in pupil dilation within this a priori region of interest. We compared the results of this analysis with results using the mean dilation (e.g., as in Mathôt et al., 2016) and peak dilation (e.g., as in Fiedler and Glöckner, 2012), which were qualitatively very similar (see Supplementary Figures B1 and B2 for results using the other two indicators). We would like to stress that our focus was not on the time course of pupil dilation because we introduced a fixed lag between payoffs, probabilities, and participants' ability to respond (however, for completeness, we plot the timecourse of pupillary responses in the Supplementary Figure B3). While some research has studied pupil dilation shortly before a decision is made (Fiedler and Glöckner, 2012, finding that pupil dilation increases as the participant is deciding), in our setup we cannot determine a unique time point at which participants made their choice: They could have reached a decision prior to being able to enter it on the rating scale.

Results

As in Experiment 1, we excluded trials in which participants indicated prices that exceeded the payoff offered in the gamble by more than $100E\$$, as this suggests lack of attention to the task. Moreover, we removed trials in which RTs deviated by $+/-3SD$ from an individual's median, assuming that these RTs are unlikely to reflect cognitive processing on a specific trial. Overall, these criteria resulted in the removal of 9.2% ($1,027/11,160$) of trials across all participants. We analyzed our data with and without these excluded trials and obtained qualitatively very similar results. We note when these exclusions led to qualitatively different results.

Pricing

As in Experiment 1, we expressed prices as deviations from a gambles' expected value ($gamble_{price} - gamble_{EV}$), thus 0 indicates perfect alignment between prices and the gambles' expected values. Across all trials, prices deviated positively from expected values in the uncorrelated condition ($b_{unc} = 76.90E\$$, $CI = [12.03E\$, 140.71E\$]$), but not the correlated conditions ($b_{neg} = -32.83E\$$, $CI = [-121.67E\$, 56.75E\$]$, $b_{pos} = 17.29E\$$, $CI = [-72.91E\$, 108.62E\$]$).[8]

As in Experiment 1, we analyzed to what extent prices differed as a function of surprise. Across all gamble types, surprising options were priced slightly higher than expected gambles ($b_{deviation} = 29.05E\$$, $CI = [-0.87; 57.44]$). This difference did not persist when comparing prices of surprising gambles to prices for reference gambles (all CIs include 0). As Figure 5 shows, pricing differences between expected and surprising options were driven by high-payoff gambles (red vs. blue bars in panels B, D).

[8] When all trials were included in the analysis, the positive deviation from expected values was more pronounced in the uncorrelated condition: $b_{unc} = 124.74E\$$, $CI = [56.51E\$, 193.98E\$]$, and slightly less so in the negative condition, $b_{neg} = -55.00E\$$, $CI = [-151.22E\$, -39.72E\$]$, $b_{pos} = -10.51E\$$, $CI = [-108.65E\$, 85.84E\$]$.

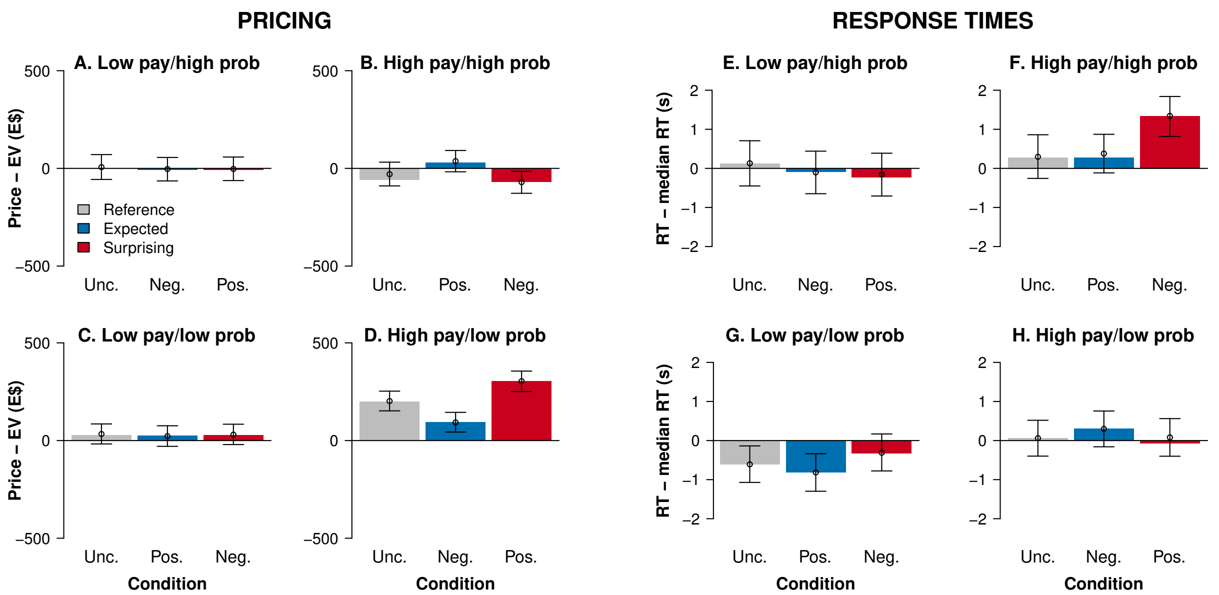


Figure 5. Behavioral results for the test gambles interspersed in Experiment 2. (A–D) Pricing. For low payoffs, there were little to no differences in how well prices of surprising gambles were aligned with the expected values of the gambles. For high payoffs, prices of surprising options differed from prices of expected options. (E–H) Response times. Participants slowed down when options were surprising. This effect was driven by the surprisingly good options (panel F). RTs expressed as deviations from individual median RTs across trials ($RT_{trial} - RT_{ind.Md}$). Black dots and error bars represent the mean and the 95% credible interval of the posterior predictive distribution.

Specifically, prices were lower for surprising high payoff/high probability gambles ($b_{deviation} = -92.78E\$$, $CI = [-172.98E\$, -12.44E\$]$). Conversely, prices were higher for surprising high payoff/low probability gambles ($b = 213.99E\$$, $CI = [42.91E\$, 385.49E\$]$, both using expected low payoff/low probability gambles as a baseline). As Figure 5 also shows, these pattern for high payoff/low probability gambles was similar when comparing suprising to reference gambles (panel D, red vs. grey bar). Taken together, these pricing patterns are consistent with the probabilities that participants could expect from high payoffs: the negative condition prompts the expectation that a high payoff will be accompanied by a low probability; the positive condition prompts the expectation that a high payoff will be accompanied by a high probability. However, prices did not seem to be adjusted to participants' expectations in their pricing of low-payoff gambles.[9] These results differ from those in Experiment 1, where prices did not deviate depending on whether a gamble had been surprising or not. One reason for these differences could be the limited presentation duration of the gambles in Experiment 2. Having less time to process the gamble while it was on screen ($4s$ in total) may have led participants to partially price the gambles based on expectations.

Response times

Participants took $4.18s$ on average per trial from seeing the empty screen ($CI = [4.60s, 4.77s]$, RT model with the uncorrelated condition as baseline) (2s enforced) to entering a response on the rating scale. There were no credible differences in these average RTs across conditions.

[9]When all trials were included in the analysis, the effect of surprise on prices high payoff/high probability gambles was not credible, but in the same direction ($b = -42.5E\$$, $CI = [-152.5E\$, 67.5E\$]$). The effect on the high-payoff, low-probability gambles was qualitatively the same ($b = 250E\$$, $CI = [45E\$, -50E\$]$).

The range of RTs was smaller in Experiment 2 than in Experiment 1 due to the experimental design, in which a price could not be entered before seeing payoff and probability information for $2s$ each. However, as in Experiment 1, our focus was on RT differences for particular gamble types. Again, RTs varied depending on the payoff–probability combination: Consistent with Experiment 1, participants across all conditions responded faster to low payoff/low probability gambles ($b = -0.85s$, $CI = [-1.37s, -0.34s]$) and spent more time evaluating high payoff/high probability gambles ($b = 1.24s$, $CI = [0.76s, 1.73s]$).

Moreover, RTs again depended on the gamble type: In Experiment 2, participants spent $0.28s$ longer on surprising gambles than on expected gambles (red. vs. blue bars, $CI = [0.030s, 0.53s]$); and $0.47s$ longer on surprising vs. reference gambles (red vs. grey bars, $CI = [0.08s, 0.88s]$). As in Experiment 1, RTs for reference gambles were statistically indistinguishable from those for expected gambles (blue vs. grey bars). As Figure 5 shows, the effect of surprise was driven by high payoff/high probability gambles gambles (panel F). The same pattern emerged when comparing surprising to reference gambles (red vs. grey bars in panels F and G; also see Supplementary Table B10). These results were qualitatively the same when data from all participants were used in the analysis.

In sum, comparing surprising (red bars) to expected (blue bars) or reference (grey bars) options in Figure 5 across panels suggests that response times for surprising options deviated most for high payoff/high probability options; that is, if the surprising option had a high expected value. It is plausible that participants were surprised by the range of possible expected values: High payoff/high probability gambles in were surprising in the negative condition (all other EVs below $600E\$$), in which such high EVs had not been experienced previously.

Pupil dilation

We analyzed pupil responses at two time points. Pupillary responses to payoffs only (t_1) were used to test the influence of payoff magnitude: If participants responded to surprising options (defined as payoff–probability combinations), pupil size should not vary between high and low payoffs at t_1. This was indeed the case across all conditions (all CIs contained 0).[10] We expected pupil dilation to change as a function of whether or not a gamble had a surprising payoff-probability combination. Figure 6 shows mean changes in pupil dilation after both payoff and probability information was presented (t_2). Panels A–D suggest that pupil size was not associated with gambles being either surprising or expected ($b = -0.14$, $CI = [-0.92, 0.63]$, main effect of surprising vs. expected across gamble types), nor with differences between expected and reference gambles ($b = 0.06$, $CI = [-1.01, 1.12]$, main effect of reference vs. expected across gamble types). Instead, there was an interaction between surprise and the gambles' expected value. Specifically, pupil size decreased in response to surprising low payoff/low probability gambles (Panel C, $b = -1.75$, $CI = [-3.38, -0.11]$) and increased in response to surprising high payoff/high probability gambles (Panel D, $b = 2.84$, $CI = [0.75, 4.96]$). When these gambles were expected or when they were reference gambles (i.e., in the uncorrelated condition), pupil responses were statistically indistinguishable from 0 (all CIs contained 0). Panels A and D show that the change in pupil dilation was statistically indistinguish-

[10] $b_{unc} = -0.07$, $CI = [-0.91, 0.76]$, using the uncorrelated condition as baseline; $b_{neg} = -0.40$, $CI = [-1.59, 0.71]$, $b_{pos} = -0.33$, $CI = [-1.52, 0.88]$, all comparisons high payoff > low payoff. These results are plotted in the Supplementary Materials. The results were qualitatively the same with and without excluded trials.

able from a 0% change for gamble types involving some tradeoff between payoffs and probabilities (high payoff/low probability or low payoff/high probability) across all three conditions—that is, irrespective of whether these gambles were surprising or not.

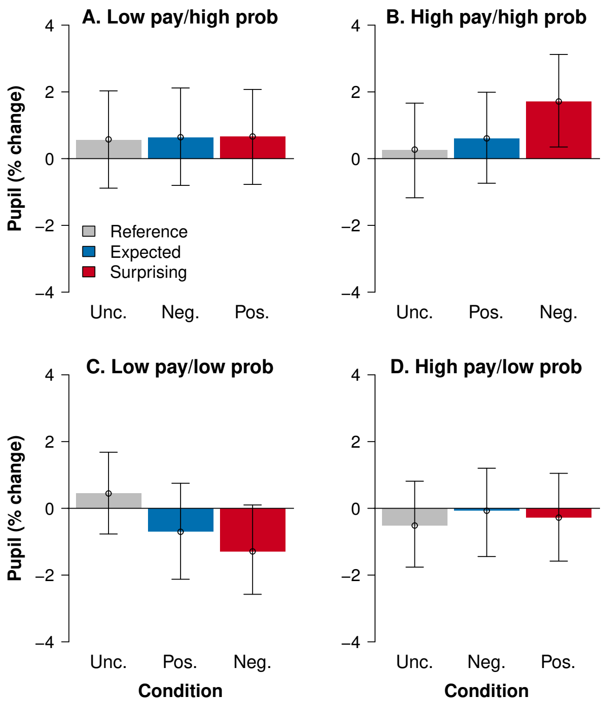


Figure 6. Pupil dilation after gambles were presented (with both payoff and probability information on the screen). Pupil dilation was computed as the median percentage change from $1 - 2s$ after the stimulus. Black dots and error bars represent the mean and the 95% posterior predictive distribution.

Summary of Experiment 2

Experiment 2 corroborates the finding that people develop expectations about the options appearing in a particular risk–reward environment. As in Experiment 1, this was evident in longer RTs for surprising gambles. This effect was driven by high payoff/high probability gambles offering larger EVs than the other gambles in the set. Moreover, there were some differences in pricing towards participants' expectations in a given environment, but this result was not robust across payoff magnitudes or across baseline comparisons. Lastly, surprising high–payoff/high–probability options were associated with a reliable increase in pupil size. However, pupil size was not associated with surprise per se.

General Discussion

If something sounds too good to be true, it usually is. The only way to identify such *exceptional* options is being sensitive to some of the tradeoffs people typically face in their choice environments. A frequent regularity people encounter is an inverse relationship between risks and rewards (Pleskac and Hertwig, 2014). We found evidence that people show distinct responses to options that deviate from the regularities and are thus surprising. That is, people seem to compare options against the risk–reward relationship they have learned from the current environment. People learned about risk–reward structures by pricing monetary gambles drawn from negative, positive, or uncorrelated risk–reward environments. After some exposure to a particular risk–reward environments, participants encountered surprising options that deviated from learned risk-reward structures. We investigated to what extent these surprising options were linked to three indicators of surprise: prices, RTs, and pupil dilation. Next, we discuss these three indicators in detail. We then relate our findings to previous research on surprise and consider their broader implications for preferential choice.

Do people *evaluate* surprising options differently from nonsurprising options? Across experiments and gamble types, we did not find robust evidence for this as indicated by the selling prices participants provided. Experiment 2 suggests that, if anything, prices deviated from gambles' expected values in the direction of participants' expectations: When the probability of surprising gambles was low but participants expected it to be high, prices were drawn towards that expectation (and were slightly higher). Conversely, when the probability was high but participants expected it to be low, prices were drawn towards that expectation (and were slightly lower). This systematicity speaks against the hypothesis that surprising options lead to more error in indicating prices that reflect gambles' EVs.

How do people *process* surprising options? Our data show that people may take more time when evaluating surprising options than when the same options are to be expected. How much longer people take to evaluate a surprising option also depends on whether the option is surprisingly good (a high payoff/high probability option) or surprisingly bad (a low payoff/low probability option). Specifically, high payoff/high probability options were surprising in the negative condition, which was characterized by otherwise similar and smaller expected values (all below $700E\$$), suggesting that "EV surprise" may amplify "risk–reward surprise" by changing participants' perceptions of possible rewards in a given experiment (also see Ludvig et al., 2014, 2018; Parducci, 1965). This "EV surprise" was also present in pupil size, which increased in response to surprisingly good options (relative to baseline), but not in response to surprise in general, or in response to high expected value in general. The differences between surprisingly good and surprisingly bad gambles are consistent with a "hold your horses" mechanism by which people behave less impulsively in the presence of high–value options (Cavanagh et al., 2014; Frank et al., 2007). That is, more scrutiny is required only when high payoffs are at stake. Here, we have shown that such a mechanism can be weakened when high payoffs are to be expected in an environment, and strengthened when high payoffs are not to be expected.

More generally, our data bring a new perspective to the growing body of research on how the environmental distribution of monetary payoffs and probabilities influences how (otherwise identical) options are evaluated (Birnbaum, 1992; Ludvig et al., 2014, 2018; Stewart et al., 2006, 2015; Walasek and Stewart, 2015). Here we show that people go beyond evaluating options from the given information. Instead, our results point to a mixture between evaluations from givens and evaluations from environments. For instance, prices for surprising options were adjusted to the options' expected values but sometimes shifted in the direction of environmental expectations. What is more, we observed systematic shifts in the way that surprising options were processed. Such environment-based evaluations cannot be anticipated by prominent theories of choice, which conceptualize risks and rewards as independent attributes that determine the expected utility of an option (von Neumann and Morgenstern, 1944) or its subjective worth (Tversky and Kahneman, 1992).

In addition, our results bring a novel dimension to studies of surprise. Similar to oddball paradigms using lower-level auditory stimuli, we show that feedback is not a prerequisite for the mind to detect surprising stimuli (or deviants) in a sequence of standards. As mentioned before, a crucial difference between perceptual oddballs and what we might call *preferential* oddballs is that the latter entails value-based stimuli. Our data suggest that, in contrast to oddball paradigms, the direction of the surprise (i.e., whether an option is surprisingly good or surprisingly bad) matters in preferential choice. Another critical difference is that attention to the stimuli is not required in oddball paradigms using auditory stimuli (Garrido et al., 2009), but is required to evaluate risky gambles. By extension, a "mismatch negativity" may not necessarily emerge for preferential oddballs—a prediction that is directly testable in further research. From this perspective, the results may be more comparable to previous studies on "risk-prediction errors" (Bossaerts, 2010; O'Neill and Schultz, 2013; Preuschoff et al., 2008) or "reward-prediction errors" (Schultz, 2002; Schultz et al., 1997) than oddball paradigms. However, these prediction errors have to date only been elicited in contexts in which feedback is given. Is feedback required for prediction errors to be elicited? This question could be investigated by comparing neurophysiological correlates of surprisingly good versus surprisingly bad options with and without explicit feedback.

A final question pertains to the role of surprise in risk–reward environments. Is there an adaptive aspect to surprise in risk–reward environments? Is its role similar to, or distinct from, other types of surprise? In value-based choice, "reward-prediction errors" are thought to aid reward-based learning. A similar mechanism would be plausible if participants anticipated the surprisingly good option being added to their bonus at the end of the study (without feedback after each choice). Similarly, in oddball paradigms, the mismatch negativity has been described as a marker for error detection that is triggered when a learned regularity is violated. It has been hypothesized that the prediction error leads to a subsequent updating of the [previously learned] model (model adjustment hypothesis, Garrido et al., 2009). This hypothesis is consistent with the predictive coding framework (Friston and Kiebel, 2009). Adjusting one's model of the world in this way could also be adaptive in risk–reward environments in which the relationships may vary (e.g., a weaker relationship is expected in newly forming markets). However, it remains an open question how much evidence is needed before a model is adjusted. Our previous experiments suggest that the risk–reward regularity people extract largely reflects the structure of the environment trials (Leuker et al., 2018). Maintaining a risk–reward rule even after seeing surprising risk–reward combinations would be consistent with a "rule-plus-exception model," according to which people may learn exceptions to a rule instead of updating that rule if they are unable to identify the rule that would account for the exceptions (Nosofsky et al., 1994). For instance, a person "might regard a single-dimension rule as tentatively acceptable as long as it correctly classifies 60% of the incoming exemplars" (p. 56).

In value-based decision making, extracting a rule from the overall environment and forming expectations may serve a very particular function: to help people to identify when options are too good to be true. Forming expectations in this way may work in many nonlaboratory environments in which risks and rewards are inversely related (Pleskac and Hertwig, 2014). In these environments, people know that there is usually "no free lunch," in that the larger rewards they desire occur only rarely (but if they are lucky and get a "free lunch" once, it does not mean that their model of risk–reward environments will change).

Ultimately, being sensitive to an environment's risk–reward structures and deviations from those structures should lead to adaptive decisions under risk.

Conclusion

Three main conclusions can be drawn from the experiments presented. First, people build expectations about the structures of the options in their global choice environments. When presented with options that deviate from these expectations, people slow down to evaluate them. Second, the direction of surprise matters: Options appear most surprising when they offer higher expected values than previously evaluated options in the set. Third, surprise can leak into peoples evaluations of the options. In value-based decision making, extracting a rule from the overall environment and forming expectations may help people to identify when options are too good to be true. Doing so can be adaptive in many environments in which risks and rewards are inversely related.

References

Becker, G. M., Degroot, M. H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, 9(3):226–232.

Birnbaum, M. H. (1992). Violations of monotonicity and contextual effects in choice-based certainty equivalents. *Psychological Science*, 3(5):310–314.

Bossaerts, P. (2010). Risk and risk prediction error signals in anterior insula. *Brain Structure and Function*, 214(5):645–653.

Brunswik, E. (1943). Organismic achievement and environmental probability. *Psychological Review*, 50(3):255–272.

Cavanagh, J. F., Wiecki, T. V., Kochar, A., and Frank, M. J. (2014). Eye tracking and pupillometry are indicators of dissociable latent decision processes. *Journal of Experimental Psychology: General*, 143(4):1476–1488.

Dalmaijer, E. S., Mathôt, S., and Van der Stigchel, S. (2014). PyGaze: An open-source, cross-platform toolbox for minimal-effort programming of eyetracking experiments. *Behavior Research Methods*, 46:913–921.

Derrington, A., Krauskopf, J., and Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *Journal of Physiology*, 357:241–265.

Fiedler, S. and Glöckner, A. (2012). The dynamics of decision making in risky choice: An eye-tracking analysis. *Frontiers in Psychology*, 3:335.

Fortin, J. (2017). Glitch in South Carolina lottery could mean $19.6 million in winnings. *New York Times*.

Fortin, J. (2018). Sorry, lottery winners. South Carolina won't pay, but here's your $1. *New York Times*.

Frank, M. J., Samanta, J., Moustafa, A. A., and Sherman, S. J. (2007). Hold your horses: Impulsivity, deep brain stimulation, and medication in parkinsonism. *Science*, 318(5854):1309–1312.

Friston, K. and Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521):1211–1221.

Gagl, B., Hawelka, S., and Hutzler, F. (2011). Systematic influence of gaze position on pupil size measurement: Analysis and correction. *Behavior Research Methods*, 43(4):1171–1181.

Garrido, M. I., Kilner, J. M., Stephan, K. E., and Friston, K. J. (2009). The mismatch negativity: A Review of underlying mechanisms. *Clinical Neurophysiology*, 120(3):453–463.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–511.

Harrington, R. (2017). A computer glitch on Christmas could give mistaken winners in the South Carolina lottery almost $20 million. *Business Insider*.

Herrmann, C. S. and Knight, R. T. (2001). Mechanisms of human attention: Event-related potentials and oscillations. *Neuroscience & Biobehavioral Reviews*, 25(6):465–476.

Huettel, S. A. and McCarthy, G. (2004). What is odd in the oddball task? Prefrontal cortex is activated by dynamic changes in response strategy. *Neuropsychologia*, 42(3):379–386.

Kruschke, J. (2015). *Doing bayesian data analysis: A tutorial introduction with R.* Academic Press, London, United Kingdom.

Leuker, C., Pachur, T., Hertwig, R., and Pleskac, T. J. (2018). Exploiting risk–reward structures in decision making under uncertainty. *Cognition*, 175:186–200.

Leuker, C., Pleskac, T. J., Pachur, T., and Hertwig, R. (2017). How the mind exploits risk–reward structures in decisions under risk. In Gunzelmann, G., Howes, A., Tenbrink, T., and Davelaar, E., editors, *Proceedings of the 39th Annual Conference Cognitive Science Society*, pages 2543–2548, London, UK.

Ludvig, E. A., Madan, C. R., and Spetch, M. L. (2014). Extreme Outcomes Sway Risky Decisions from Experience. *Journal of Behavioral Decision Making*, 27(2):146–156.

Ludvig, E. A., McMillan, N., Madam, C. R., Xu, H., and Spetch, M. L. (2018). Living near the edge: how extreme outcome and their neighbours drive risky choice. *Journal of Experimental Psychology: General*.

MacLeod, D. I. A. and Boynton, R. M. (1979). Chromaticity diagram showing cone excitation by stimuli of equal luminance. *Journal of the Optical Society of America*, 69(8):1183–1186.

Mathôt, S., van der Linden, L., Grainger, J., and Vitu, F. (2016). The pupillary light response reflects eye-movement preparation. *Journal of Experimental Psychology: Human Perception and Performance*, 41(1):28–35.

Näätänen, R. (2007). The mismatch negativity: Where is the big fish? *Journal of Psychophysiology*, 21(3–4):133–137.

Nosofsky, R. M., Palmeri, T. J., and McKinley, S. C. (1994). Rule-plus-exception model of classification learning. 101(1):53–79.

O'Neill, M. and Schultz, W. (2013). Risk prediction error coding in orbitofrontal neurons. *The Journal of Neuroscience*, 33(40):15810–15814.

Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review*, 72(6):407–418.

Peirce, J. W. (2007). PsychoPy - Psychophysics software in python. *Journal of Neuroscience Methods*, 162(1–2):8–13.

Picton, T. W. (1992). The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology*, 9(4):456–479.

Pleskac, T. J. and Hertwig, R. (2014). Ecologically rational choice and the structure of the environment. *Journal of Experimental Psychology: General*, 143(5):2000–2019.

Preuschoff, K., Quartz, S. R., and Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *Journal of Neuroscience*, 28(11):2745–2752.

Preuschoff, K., 't Hart, B. M., and Einhäuser, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline's role in decision making. *Frontiers in Neuroscience*, 5:115.

Rüsseler, J. and Rösler, F. (2000). Implicit and explicit learning of event sequences: Evidence for distinct coding of perceptual and motor representations. *Acta Psychologica*, 104(1):45–67.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron*, 36(2):241–263.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63(2):129–138.

Skylark, W. J. and Prabhu-Naik, S. (2018). A new test of the risk-reward heuristic. *Judgment and Decision Making*, 13(1):73.

Squires, N. K., Squires, K. C., and Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, 38(4):387–401.

Stan Development Team (2016). RStanArm: Bayesian applied regression modeling via Stan (version 2.9.0-4). R package version 2.13.1.

Stewart, N., Chater, N., and Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, 53(1):1–26.

Stewart, N., Reimers, S., and Harris, A. J. (2015). On the origin of utility, weighting, and discounting functions: How they get their shapes and how to change their shapes. *Management Science*, 61(3):687–705.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323.

Van Steenbergen, H. and Band, G. (2013). Pupil dilation in the Simon task as a marker of conflict processing. *Frontiers in Human Neuroscience*, 7:215.

von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ.

Walasek, L. and Stewart, N. (2015). How to make loss aversion disappear and reverse: Tests of the decision by sampling origin of loss aversion. *Journal of Experimental Psychology: General*, 144(1):7–11.

4 | Risk–reward structures can promote satisficing in decisions under risk

Leuker, C., Pachur, T., Hertwig, R. & Pleskac, T.J. *In Prep.*

Abstract

Risk–reward structures can vary across environments. While they can be relatively uncorrelated in newly forming markets, they are inversely related in many monetary and nonmonetary domains. As risks and rewards have by and large been treated as uncorrelated in empirical studies of risky choice, theories derived from them may only capture a small fraction of how people make such decisions. Here, we investigated how the relationship among risks and rewards impacts how people process options and form preferences in decisions under risk. In our experiment, participants learned about risk–reward structures being either negatively or positively correlated, or uncorrelated, from pricing gambles drawn from the environments (incidental learning). In a test phase, participants chose among two nondominated gambles under different amounts of time pressure and while their eye movements were tracked. Behavioral data and computational modeling with a drift-diffusion model suggested little influence of risk–reward environments on preferences. However, the risk–reward environment affected how gambles were processed. Specifically, exposure to an environment where risks and rewards were uncorrelated resulted in a more even distribution of gaze and more cautious responding compared to exposure to an environment in which risks and rewards were inversely related, where participants satisficed sooner. These results did not generalize to (1) conditions where participants learned a risk–reward structure explicitly from feedback, (2) conditions with very strong time pressure (1.5s response deadline), (3) learning about a counter-intuitive, positive risk–reward relationship. In sum, uncorrelated risk–reward environments can promote maximizing—that is, more rigorous processing and closer-to-normative choices; while environments in which risks and rewards are inversely related allow, and tend to promote, satisficing.

This experiment was preregistered on the Open Science Framework. | Acknowledgments: We would like to thank Lasare Samartzidis for assistance with data collection.

Introduction

According to most accounts, the methods used to study risky decision making can be largely traced back to the curiosity of one person: Chevalier de Méré. This so-called gentleman gambler of France was curious why he lost money on certain bets and in other cases he was just curious about gambling puzzles he had heard about, perhaps at the salon. He often turned to his friend Blaise Pascal to help appease his curiosity. One of these puzzles was the problem of points, and in solving this problem Pascal, along with Pierre de Fermat, laid the groundwork for what would become the predominant theory of probability (David, 1962; Daston, 1988; Hacking, 1975; Todhunter, 1865).[1]

Pascal's solution, however, had another effect. As George Boole (1854) noted, the problem of points was "... the first of a long series of problems, destined to call into existence new methods in mathematical analysis, and to render valuable service in the practical concerns of life" (p. 243). The problems were, of course, gambling problems like the problem of points and the practical concerns were risk and risky decision making. Looking back, Pascal and Fermat's focus on gambling problems has proven a useful recipe for other scholars to cook up insights into how people should and do make such decisions (e.g., Allais, 1953; Bernoulli, 1954; Edwards, 1954; Ellsberg, 1961; Kahneman and Tversky, 1979; von Neumann and Morgenstern, 1947; Savage, 1954). The recipe is to design specific gambles that serve as critical tests for decision theories (Birnbaum, 2011; Pleskac et al., 2015). If choices in these gambles contradicted the predictions of the theory, the theory was modified. Take for instance the following choice problems that Maurice Allais (1953) used to show that people's choices deviate from the predictions of expected utility theory. The original formulation asked people to make two choices. The first choice was between the following two options:

$$\text{Option A}^*: \quad 100\% \text{ chance of winning 100 million Fr.}$$

$$\text{Option B}: \begin{cases} 10\% \text{ chance of winning 500 million Fr.} \\ 89\% \text{ chance of winning 100 million Fr.} \\ 1\% \text{ chance of winning 0} \end{cases}$$

Most people prefer the 100 million Fr. with certainty in Option A (*). Now consider a second choice that Allais created by removing the 89% chance of winning 100 million Fr. from both options:

[1] The problem of points refers to a game of chance with two players who have equal chances of winning each round. The players contribute equally to a prize pot, and agree in advance that the first player to have won a certain number of rounds receives the prize. The question de Méré asked was if the game is interrupted before either player has won, how should the pot be divided? The starting insight for Pascal and Fermat was that the division should not depend so much on the history of the part of the interrupted game that actually took place, as on the possible ways the game might have continued, if it had not been interrupted.

$$\text{Option C} : \begin{cases} 11\% \text{ chance of winning 100 million Fr.} \\ 89\% \text{ chance of winning } 0 \end{cases}$$

$$\text{Option D*} : \begin{cases} 10\% \text{ chance of winning 500 million Fr.} \\ 90\% \text{ chance of winning } 0 \end{cases}$$

Now most people prefer taking their chances to win 500 million Fr. (Option D). As this pattern of preferences is inconsistent with expected utility theory, it has been used to question the generality of using expected utility theory as a descriptive theory of choice (e.g., MacCrimmon and Larsson, 1979). This set of problems, in fact, served as the foundation for the development of prospect theory as a descriptive theory of risky choice (Kahneman and Tversky, 1979).

However, to people who are not focused on testing decision-making theories, there is something quite impractical with these problems: the payoffs are quite large. Most of us do not get to make choices among millions of francs (or euros, or dollars) very often. Allais' gambles seem particularly extreme: The approximate value of 500 million French Franc from 1950 is €10 million in 2018. Moreover, these millions of Francs were linked to extremely high chances of winning (89% and 90% in the second choice problem). The use of gambles offering large, likely payoffs has not just been limited to Allais. Consider the choice between 4000 with a probability of 80% or 3000 for sure. (Kahneman and Tversky, 1979) used this gambling problem and another problem where the same people were offered a choice between 4000 with a probability of 20% or 3000 with a probability of 25% as a simpler version of the problems Allais used. But, even here the magnitudes of the hypothetical payoffs referred to Israeli pounds and were quite large. According to Kahneman and Tversky (1979) the "outcomes [used in the choice problem] refer to Israeli currency. To appreciate the significance of the amounts involved, note that the median net monthly income for a family is about 3,000 Israeli pounds" (p. 264). These values and the associated probabilities are clearly less extreme than those used by Allais (1953), but still rather unrepresentative of gambles people are typically offered.

The top left panel of Figure 1 plots all the gambles that Kahneman and Tversky (1979) used. It shows that this issue of unrepresentative gambles is not restricted to a few specific gambles but largely pertains to all the gambles they used to motivate the development of prospect theory. Plotting the gambles this way also reveals a potentially larger issue: Across the gambles there is little to no relationship between the probabilities and payoffs. The remaining panels of Figure 1 demonstrate that across seven other prominent papers on risky choice there is a similar pattern in that any given payoff can occur with almost any probability. That is, risks and rewards are uncorrelated. In some situations outside the laboratory, risks and rewards *can be* uncorrelated—for instance if a market has not yet reached an equilibrium, if there are more resources than competitors in an environment or if competitors have imperfect information on where to seek for rewards (Pleskac et al., prep). However, frequently and across many monetary and nonmonetary domains, risks and rewards are inversely related (Pleskac and Hertwig, 2014). Therefore, theories derived from choice environments with uncorrelated risks and rewards may have little correspondence to the choices

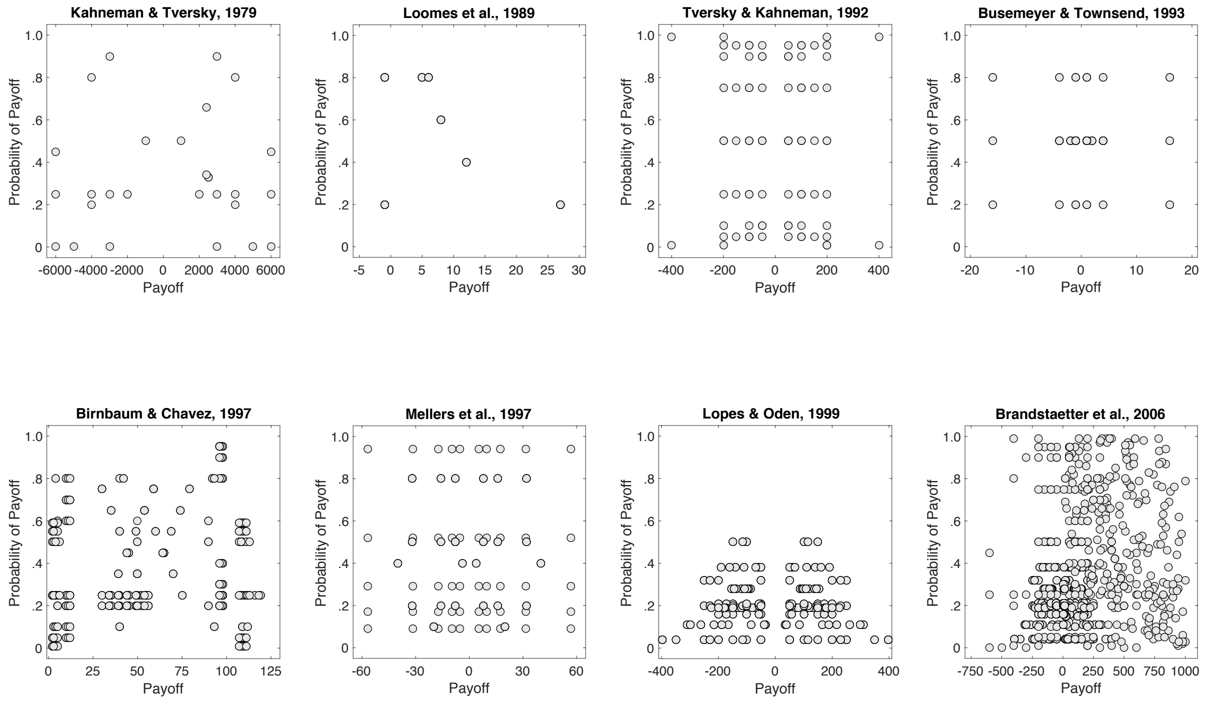


Figure 1. Relationship between payoffs and probabilities realized in laboratory investigations of prominent theories of risky choice. Figure reproduced from Pleskac and Hertwig (2014).

people make in nonlaboratory environments; or at the very least they capture only a small fraction of choice situations people typically face. As Brunswik stated, not acknowledging such environmental regularities results in experimental manipulations that are "more like a mere homunculus of the laboratory out in the blank" (1955, p. 204). In other words, many studies on risky choice may have compromised external validity to increase internal validity (Dhami et al., 2004), in this case by being able to characterize choices in the entire set of theoretically conceivable risk–reward relationships.

One reason nonrepresentative gambles in risky choice experiments can create a problem follows from theories of adaptive cognition. According to these theories, a mind that has learned the relationship between key variables in the environment—such as the frequent and recurrent relationship between risks and rewards—can subsequently exploit the relationship between these variables (Brunswik, 1952; Gigerenzer et al., 1991; Gibson, 1979; Marr, 1982; Simon, 1956; Stewart et al., 2006). For instance, people can decide to rely on a subset of cues in the environment because cues are interrelated (Brunswik, 1952). People, in fact, exploit risk–reward structures in decisions under uncertainty—in form of a risk–reward heuristic—to infer probabilities directly from the magnitude of the payoffs (Pleskac and Hertwig, 2014; Leuker et al., 2018a), or to infer payoffs from probabilities (Skylark and Prabhu-Naik, 2018). We were curious whether a similar mechanism may be in place in decisions under risk—where the probabilities of a payoff are, in principle, available. In correlated risk–reward environments, one of the attributes is redundant and can, in principle, be inferred from the other. This may result in simpler, less rigorous processing strategies. Conversely, such processing strategies may be suspended if risks and rewards are uncorrelated.

We are not the first to investigate the how the structure of the environment impacts risky choice. One structure that has been investigated is dominance among gambles. One gamble A dominates another gamble B when it has both a higher payoff ($x_A > x_B$) and a higher probability ($p_A > p_B$). In such environments, simulations have shown that more frugal, heuristic strategies (e.g., only considering payoffs or only considering probabilities before making a choice) and more rigorous, compensatory strategies (that seek to maximize EV by first muliplicatively combining payoffs and probabilities within each gamble, $EV_i = p_i \times x_i$) perform equally well. The opposite pattern holds when gambles are non-dominated (i.e., gamble A offers a higher payoff x, but gamble B offers a higher probability p: $x_A > x_B$ and $p_A < p_B$). In nondominated choice environments, more rigorous, compensatory choice strategies perform better (Payne et al., 1988).

Interestingly, nondominated gambles actually have a risk–reward structure in that by definition the gamble with the higher payoff will have a lower probability and vice versa. Thus, our Brunswikian prediction—that correlated risk–reward structures produce more frugal processing strategies—seems to be at odds with the simulations from Payne et al. (1988), that imply more rigorous processing is to be expected or at least preferable when gambles are nondominated. However, in the simulations and in most empirical studies of risky choice, the dominance relationship was only created between the two gambles a (simulated) person had to choose between at a given time. Consequently, the risk–reward structure is restricted to the microcosm between the two nondominated gambles. That is, nondominated options create a *local* risk–reward structure. Between choice problems, there can still be great latitude in the way in which payoffs and the probabilities are combined: none of the gambles in Allais' choice problems were dominated and yet these options do not fall within the range of options one typically encounters. This implies that the global risk–reward structure or the correlation between probabilities and payoffs across all possible gambles may be an important context variable. Environments in which risks and rewards are (assumed to be) correlated may invite more frugal processing strategies—even if gamble pairs are locally nondominated.

Overview of experiment and hypotheses

We investigated how exposure to one of three different risk–reward structures (negative, positive, uncorrelated) impacted participants processing strategies and choices under different processing demands. Briefly, we created a separate learning phase in which participants learned about the different risk–reward environments. The learning phase was identical to the learning phase in Leuker et al. (2018a), in which we showed that people learn about risk–reward relationships incidentally.[2,3] The key analyses are based on the test phase in which participants chose among two nondominated options of the form "p chance of winning x, otherwise nothing". Within–participants, we manipulated whether choices had to be made under strong timepressure (within 1.5s, "fast") or whether making the best possible choice was emphasized (no timepressure, "best"). We created a separate learning and test phase to be able to demarcate between participants learning or updating their risk–reward "priors", and subsequent effects on choice. We

[2] We also created a learning phase in which participants learned about risk–reward relationships explicitly, and compare the two in the general discussion.

[3] As with Leuker et al. (2018a), the experiment included posttests in which participants were asked for their payoff–dependent probability estimates, in order to assess the risk–reward structures participants assumed to be present in the experiment. Moreover, participants chose between an uncertain option and a sure thing to replicate earlier results showing that risk–reward structures are exploited in decisions under uncertainty. These tasks and their results are reported in the Supplementary Material.

preregistered our hypotheses on the Open Science Framework after a first, exploratory experiment (Leuker et al., 2017).[4] We summarize key hypotheses and how they were operationalized below.

Processing hypotheses

Our Brunswikian prediction was that a risk–reward structure affords people the opportunity to infer the probabilities from the magnitudes of the payoffs themselves (or vice versa) and therefore invites more frugal processing of monetary gambles where people only attend to one attribute of the gambles (e.g., the payoffs)—even if gamble pairs are locally nondominated. One qualification of this prediction is that people may need a good reason to ignore information that is, in principle, available. As Payne et al. (1993) found, the extent to which people exploit statistical regularities in their environments can depend on the need to reduce processing demands. Inspired by these theories, we hypothesized that participants would rely more on frugal processing strategies when time was limited—and more so in correlated risk–reward environments than in uncorrelated ones. In other words, participants in a negative or positive risk–reward environment may shift processing strategies to be more frugal (ignoring some information), whereas participants in an uncorrelated risk–reward environment may handle timepressure differently, for example by attempting to accelerate processing and process the same information at a higher rate (Zur and Breznitz, 1981; Payne et al., 1993).

Choice hypotheses

If an uncorrelated risk–reward environment leads to more rigorous processing strategies where both payoffs and probabilities are examined, this should result in a higher proportion of expected value maximizing choices than for instance the principle of indifference that assumes equal probabilities (.5) for all options (Fox and Clemen, 2005; Fox and Rottenstreich, 2003; Thorngate, 1980). It has been shown empirically that people do not completely ignore any of the attributes in globally uncorrelated risk–reward environments, the standard environments used in most experimental settings (Stewart et al., 2006; Manohar and Husain, 2013; Fiedler and Gloeckner, 2012; Pachur et al., 2018). It could also be that people set aspiration levels, i.e. that because they want to maximize payoffs, they inspect options more closely *and* they consider that there is something to be gained from inspecting all attributes in their environment (i.e. a high marginal improvement from searching more). In other words, the more rigorously options are examined, the more closely processing strategies and resulting choices can maximize expected value (Simon, 1978).

In a perfectly negative risk–reward environment, payoffs and probabilities trade off against each other within each alternative (e.g. Do you prefer \$99 with $p = .01$ or \$1 with $p = .99$?). This means expected values—for example in a choice among two options drawn from an environment—will be identical, or almost identical if the relationship is noisy. In this case, a decision maker does not need to compute expected values carefully. Instead, she can rely on her preferences (e.g. choosing the higher–probability gamble, \$1 with $p_{inferred} = .99$). This view is consistent with prior research, suggesting that difficult choice problems are associated with more heuristic processing (Pachur et al., 2013). If the risk–reward

[4]The experiment in Leuker et al. (2017) showed that risk–reward structures for some individuals led to more frugal processing strategies under moderate timepressure (choices within $3s$), but we cannot rule out that the distribution of possible EVs produced these results. We controlled for this aspect in the experiment presented here.

relationship is somewhat noisy, this may come at the cost of choosing the lower expected value option. However, if the differences in the expected values between the two gambles are small, the overall loss may not be substantial.

A positive risk–reward environment in which high payoffs are linked to high probabilities and low payoffs are linked to low probabilities is counterintuitive. Consequently, this positive risk–reward relationship may sometimes be learned less well compared to a negative risk–reward structure (Leuker et al., 2018a). However if the structure is learned and fully exploited—for instance, by only attending to payoffs–participants should always choose the higher—payoff option and *assume* dominance of this option over the other (because the higher payoff is assumed to be linked to a higher probability; e.g. $\$99 \times p_{inferred} = .99 > \$1 \times p_{inferred} = .01$).[5] In other words, exposure to correlated environments may lead people to terminate search sooner, because more search takes time, and at the same time, more search may only lead to marginal improvement in choice. Such strategies are sometimes referred to as "satisficing" (Simon, 1978).

Methods

Participants

A total of 186 adults from the participant pool maintained at the Max Planck Institute for Human Development completed the experiment. Being unsure of effect sizes, but expecting them to be small (see preregistration), our sample size was based on our previous studies, and exceeds the sample size of earlier, comparable eyetracking studies (Gloeckner and Herbold, 2011; Stewart et al., 2016). We excluded two participants who chose the dominated alternative in three out of the five catch trials (under no timepressure), and three participants due to poor eyetracking. We analyzed data from 181 participants (107 females, mean age = 26.09, $SD = 4.94$). All participants were paid a fixed rate of €12 plus a bonus based on their performance in the learning phase and choice task (€$1.13 - 10.04$).
Half of the participants learned about risk–reward relationships incidentally as they priced gambles from one of three risk–reward environments. The other half of the participants learned the risk–reward structure explicitly, via a function learning task. We did not predict differences between these two learning conditions a priori. However, systematic differences emerged between the two learning conditions, with the function learning task appearing to bias attention to payoffs by task design (as payoffs were used as a cue across all conditions and trials without a counterbalanced version of the task). Moreover, in the explicit learning conditions there was an asymmetry between the risk–reward environments such that participants in the correlated environmnent needed substantially fewer trials to learn the risk–reward relationship as compared to participants in the uncorrelated environments who needed many trials to learn no relationship existed. This asymmetry may have frustrated some participants in the uncorrelated condition. For these reasons, we focus on the participants who learned incidentally ($N = 89$). The Supplementary Material reports the results from the explicit learning conditions. We briefly discuss the differences between the learning

[5]Originally, we hypothesized that participants in the correlated risk–reward conditions would achieve similar "accuracy" levels for gamble pairs that fit their previously learned risk–reward environment (see H2 in preregistration). However, this prediction is not testeable with "p (choose higher EV)" as a dependent variable because inferring probabilities from payoffs in the negative risk–reward environment leads to equal expected values for both gambles, which warrants either random, or risk–preference dependent choice (but is silent about a choice rule/how the higher EV choice is detected).

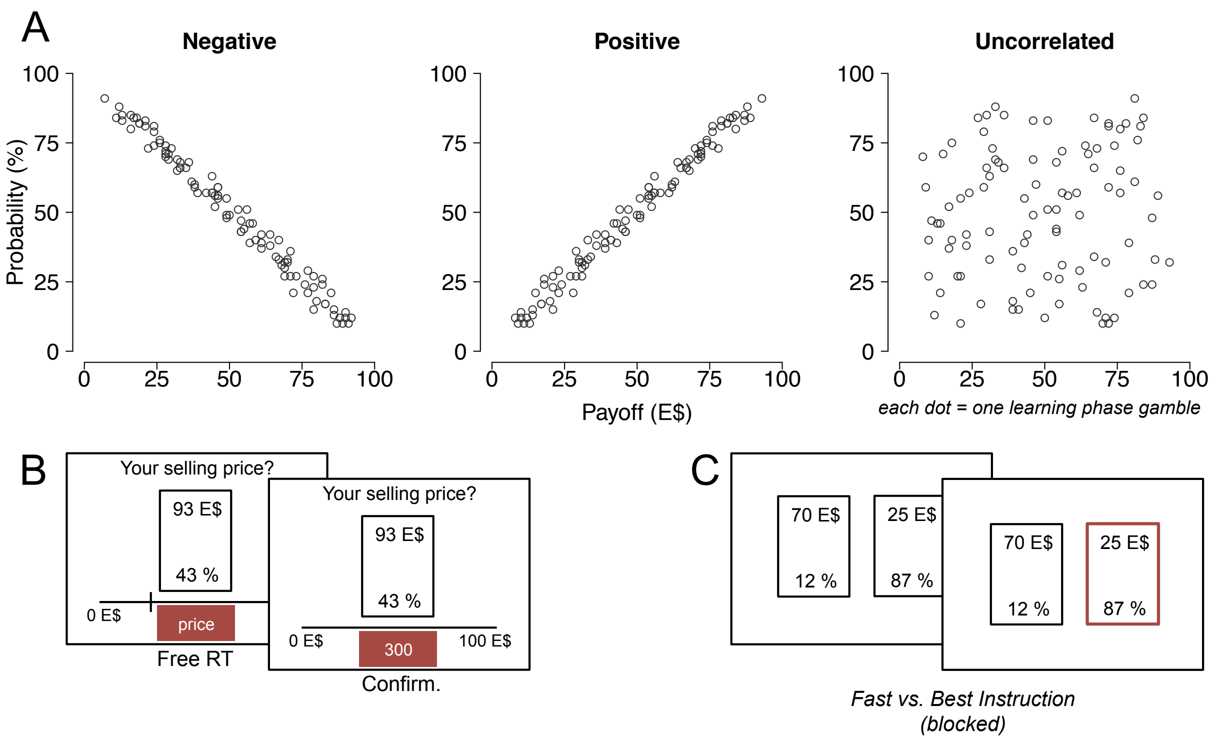


Figure 2. Experimental Setup. (A) Learning phase gambles were drawn from one of three risk–reward environments. (B) Participants learned about one of these risk–reward environments incidentally, from pricing gambles. (C) Participants completed a test phase where they chose among two nondominated gambles, while their eyemovements were being tracked. In the test phase, 60% of the gambles were common across risk–reward environments, 40% were in line with a previously learned structure. Eyetracking analyses are based on the first screen in the choice phase, before participants indicated their choice. *Best* (no timepressure) and *fast* choices were made in interleafed blocks with 16 choices each.

conditions and provide a possible account for systematic differences between them in the general discussion. All gambles were in the gain domain. The experiment was approved by the Ethics Board at the Max Planck Institute for Human Development.

Procedure

Incidental learning phase (pricing).

In the incidental learning phase, participants were presented with 100 gambles and asked to indicate their willingness to sell for each of them (Figure 2B). Between subjects, the majority of gambles were constructed such that across gambles, payoffs and probabilities were negatively correlated, positively correlated, or uncorrelated (Figure 2A). To motivate participants to report their true valuations of the gambles, we implemented a Becker-DeGroot-Marschak auction (Becker et al., 1964). In particular, ten gambles were selected at the end of the experiment and participants either played out the gamble or received their stated selling price. Participants took self-paced breaks after each of five blocks; each block consisted of 20 gambles. As participants were never asked to pay attention to the underlying risk–reward structure nor was learning the structure the central task, we refer to this learning as incidental learning (Leuker et al., 2018a). Across all gambles in this experiment, we used the same range for payoffs and probabilities $(1 - 100)$ to control for the number of digits to be expected per attribute, and thus not bias attention to either payoffs or probabilities by making one of them harder to read. Moreover, we used an experimental currency, the $E\$$ (conversion rate $100E\$ = €1$, disclosed in the instructions).

Test phase (choice).

Participants repeatedly chose between two monetary gambles of the form "p chance of winning x, otherwise nothing." In addition, participants were informed that there were two different types of rounds: Rounds in which they had to decide within $1.5s$ (fast) and rounds in which they should make the "best possible decision" (best). These rounds were presented in interleaved blocks (16 choices per block), and participants would receive information about the upcoming block before it started ("Time limit – Make a fast decision" vs. "No time limit – Make a good decision").[6] When taking longer than $1.5s$ in time pressure trials, participants lost the chance to win a bonus if that particular trial was played out at the end. Participants completed five practice trials with and five practice trials without time pressure to familiarize themselves with the task. Across participants, we randomized the positions of the gambles on screen, counterbalanced the location of payoffs and probabilities (top/bottom), and counterbalanced the order of blocks (best/fast blocks first).

The gamble pairs in the choice phase were partly drawn from the same, condition–dependent risk–reward environments as participants had been exposed to in the learning phase (proportion = .4) and partly common across the different risk–reward environments (proportion = .6). In each trial, two gambles were paired such that neither option was dominated. The same set of gambles was used in the best vs. fast trials—thus, each participant was presented twice with each gamble pair. The focal analyses are based on the common set of gambles. This analysis strategy was preregistered. Analyzing differences in processing strategies on the common gamble pairs was important in order to control for expected value differences which may otherwise have confounded the analyses.[7] Within the common gamble pairs, there were different *gamble types* (see https://osf.io/yx5m8/). Condition–dependent processing strategies did not systematically vary across these gamble types and thus we collapsed across them in our analyses. The expected value differences for the common gambles were approximately $7E\$$ (absolute difference: $Md = 7.02$, $IQR = 0.85 - 8.7$, ratio between the larger EV gamble and the smaller EV gamble: $Md = 2.01$, $IQR = 1.26 - 2.69$; also see Supplementary Figure C2).

Eye Tracking

During the choice phase, we collected binocular eye position data with an EyeTribe tracker, sampled at $60Hz$. The experiment was implemented in PsychoPy 1.83.01 and the eye-tracking interface PyTribe (Dalmaijer et al., 2013). Each participant's eye movements were calibrated using the Eyetribe UI with a 9-point grid before each task (< 0.7 degrees of visual angle). Participants were seated approximately

[6] The German wording was: "Zeitlimit - entscheide dich schnell" and "Kein Zeitlimit - entscheide dich gut"

[7] In our first experiment, the risk–reward structure was maintained in the test phase (Leuker et al., 2017). Although controlling for EV differences statistically is possible, one challenge was to disentangle the extent to which EV differences vs. risk–reward structures *induced* certain processing strategies: Drawing a random set of nondominated options from a globally uncorrelated risk–reward environment means that gamble pairs can have large variations in their EV differences. Drawing a set of nondominated options from a globally positively correlated risk–reward environment means that gamble pairs have extremely small, sometimes indiscriminable, EV differences. Thus, there is an inherent tradeoff between controlling for EV differences and maintaining global risk–reward structures in a choice task that we will come back to in the discussion.

$60cm$ from the screen using a chinrest affixated to the table, in a room with negligible ambient light. We preprocessed raw samples by parsing eye-tracking data into fixations and saccades using the saccades package in R (Von der Malsburg, Titus, 2015). Eye-tracking analyses in this paper are based on fixation data.

Analyses

We used a Bayesian approach to data analysis (Kruschke, 2014). Specifically, we applied Bayesian Generalized Linear Mixed Models using Stan in R for regression analyses with the rstanarm package (Stan Development Team, 2016). We entered participant as a grouping factor to account for individual variation beyond condition-dependent effects. We ran three chains using a Markov Chain Monte Carlo sampler to draw from posterior distributions of parameters. Depending on model complexity, we ran $10{,}000 - 30{,}000$ samples per chain to ensure an effective sample size of 10,000 for each coefficient and set a burn-in of 500 samples. We investigated (convergence of) our posteriors through visual inspection and the Gelman–Rubin statistic (Gelman and Rubin, 1992). We report the mean of the posterior distribution of the parameter or statistic of interest and two-sided 95% equal tail credible intervals (CI) around each value. Our focus is on estimating the effects of particular conditions and our analyses reflect this goal; in comparing the conditions, however, the crucial issue was whether the credible values included 0 or not.

Results

Behavioral (learning phase)

The prices participants set for the gambles were well adjusted to expected values and this was consistent across all three risk–reward conditions. There were no reliable differences between conditions. On average, prices exceeded the corresponding expected value of the gamble by around $7E\$$ ($M_{deviation} = 7.12E\$$, $b_{neg.} = 9.6E\$$, CI = $[6.7; 12.5E\$]$; $b_{pos.>neg.} = -4.0E\$$, CI = $[-8.1, 1.0E\$]$; $b_{unc.>neg.} = -3.5E\$$, CI = $[-7.6, 4.8E\$]$).

Behavioral (test phase)

Participants appeared to be well-adjusted to the time pressure manipulation with only 306 trials having responses outside of the response window (306/23552 trials across all participants, proportion = .04). There was no credible difference in the number trials that fell outside of the response window between conditions ($b_{pos.>neg.} = 0.21$, CI = $[-0.29, 0.71]$; $b_{unc.>neg.} = -0.15$, CI = $[-0.34, 0.69]$).

Figure 3A shows that in all conditions, the higher EV option was chosen above chance level (.5). Modeling choices on a trial-by-trial level revealed that larger expected value differences were more likely to lead to EV maximizing choices across conditions ($b = 0.084$, CI = $[0.078, 0.091]$). As the right panel in Figure 3A shows, and as expected, the fast instruction strongly decreased how often the higher expected value option was chosen ($b = -0.69$, CI = $[-0.81, -0.57]$); models using participant ID as a grouping factor and the negative condition as a baseline). The effect of timepressure was similar across risk–reward conditions ($b_{pos.>neg.} = -0.010$, CI = $[-0.19, 0.17]$; $b_{unc.>neg.} = -0.12$, CI = $[-0.30, 0.06]$). Following our

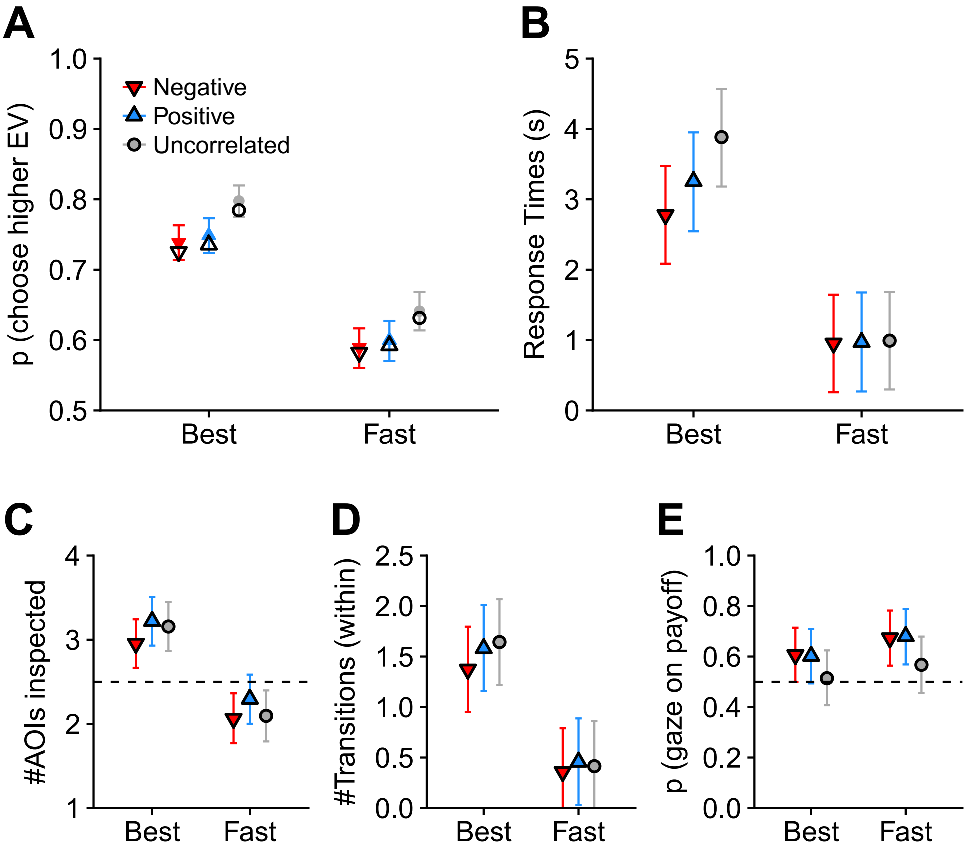


Figure 3. Descriptive results (test phase). Colors indicate means and 95% HDIs of the posterior predictive distributions. Black triangles/circles indicate means. Posterior predictive distributions are based on a model accounting for individual variation and EV differences. Posterior predictive distributions in A are based on the median EV difference in the experiment (7*E*\$), and across EV differences in the other panels. Dashed lines in C indicate the average number of AOIs that can be inspected (2.5). Dashed line in E indicates an equal distribution of gaze to both payoffs and probabilities.

hypotheses, we then tested how learning about the different risk–reward environments impacted choice. To do so, we modeled whether participants chose the option with the higher expected value using risk–reward condition as a predictor, while controlling for expected value differences and individual variation. This analysis revealed that participants in the uncorrelated condition were approximately 1.4 times more likely to pick the higher expected value option in a given trial in the best instruction ($M_{neg.} = .71$; $M_{unc.} = .76$, $b_{unc.>neg.} = 0.35$, CI $= [0.11, 0.59]$). Participants in the positive conditions also maximized expected values more often than participants in the negative condition, but the credible interval included 0 ($M_{pos.} = .75$, $b_{pos.>neg.} = 0.051$, CI $= [-0.03, 0.29]$).

Figure 3B depicts the average response times in the three risk–reward conditions. When instructed to make the best possible choice (left panel), participants in the negative condition responded sooner ($M_{neg.} = 2.78s$) than participants in the other two conditions ($M_{pos.} = 3.26s$, $b_{pos.>neg.} = 1.11$, CI $=$ $[0.32, 1.91s]$; $M_{unc.} = 3.88s$, $b_{unc.>neg.} = 0.85s$, CI $= [0.05, 1.66s]$). In the fast instruction, there were no credible differences in response times between conditions ($b_{pos.>neg.} = 0.023$, CI $= [-0.052, 0.098]$; $b_{unc.>neg.} = 0.038$, CI $= [-0.034, 0.098]$).

Processing strategies (test phase)

As a first approximation of how rigorously participants processed the options before making a choice, we computed the number areas of interest (AOIs) a participant inspected in each trial. Inspecting more AOIs might suggest more rigorous processing whereas inspecting only a subset of AOIs might suggest more frugal processing. The number of AOIs inspected can range from 1 to 4, as on each screen participants were presented with two payoffs and two probabilities. Figure 3C plots the average number of AOIs inspected by risk–reward condition.

A comparison between the left and the right panel suggests that, the main reason to ignore some of the attributes was the need to make fast choices. Under timepressure, participants in the negative condition inspected 2.06 attributes on average (CI = $[1.78, 2.33]$), participants in the other two conditions inspected slightly more attributes, but these differences were small—and the credible interval included 0 ($b_{pos.>neg.} = 0.23$, CI = $[-0.15, 0.62]$); $b_{unc.>neg.} = 0.02$, CI = $[-0.36, 0.39]$, after controlling for EV differences and individual variation).

When instructed to make the best possible choice, a similar pattern emerged between risk–reward conditions. Participants in the negative condition inspected 2.98 attributes on average (CI = $[2.74, 3.21]$), whereas participants in the other two conditions seemed to sample information more carefully ($b_{pos.>neg.} = 0.28$, CI = $[-0.06, 0.60]$); $b_{unc.>neg.} = 0.20$, CI = $[-0.12, 0.53]$; after controlling for individual variation and expected value differences). Again, these differences were small—and the credible interval included 0. If participants in the negative condition inspected fewer attributes because they inferred one attribute from the other, there should be fewer within–gamble transitions in the negative condition. While the direction of the estimated differences between conditions speaks for such a systematicity, the differences were not reliable ($b_{pos.>neg.} = 0.23$, CI = $[-0.10, 0.54]$); $b_{unc.>neg.} = 0.29$, CI = $[-0.02, 0.60]$; plotted in Figure 3D). There were reliable differences in how participants distributed their gaze across the options, with the uncorrelated condition distributing their gaze more evenly across payoffs and probabilities (i.e, proportion of gaze on payoffs: $M = 0.51$, CI = $[0.39, 0.65]$), whereas gaze in the negative condition was reliably different from .5, with a bias towards payoffs ($M = 0.60$, CI = $[0.55, 0.66]$). The difference in gaze patterns in the uncorrelated vs. negative condition was credible ($b_{unc.>neg.} = -0.09$, CI = $[-0.16, -0.01]$; plotted in Figure 3E).

Processing strategies and choices (test phase)

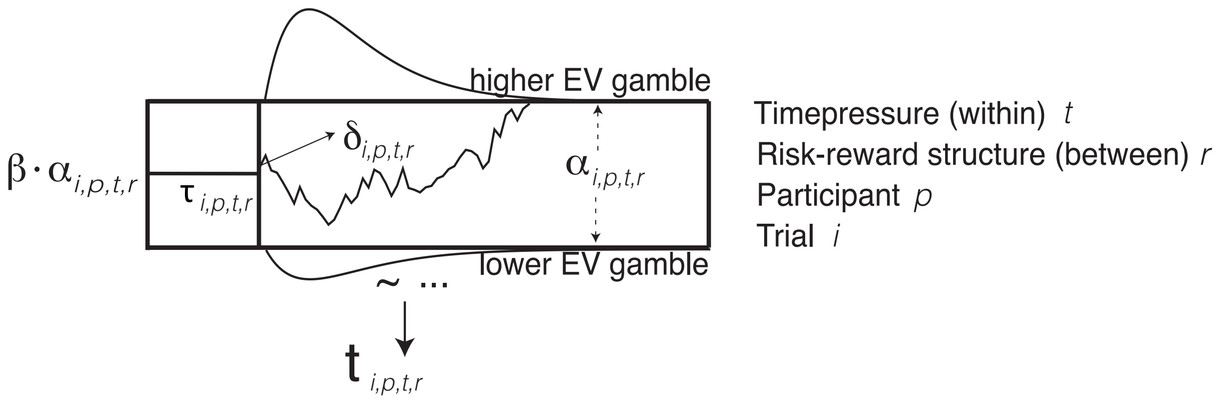
Lastly, we assessed whether more rigorous processing strategies were linked to choosing the higher expected value option on a given trial. Consistent with prior research (Payne et al., 1988), the higher expected value option was chosen more on trials in which more attributes were inspected ($b_{fast} = 0.17$, CI = $[0.11, 0.23]$, $b_{best} = 0.12$, CI = $[0.05, 0.18]$, across environments after controlling for EV differences and individual variation). This is a direct consequence of the environment consisting of nondominated option pairs, in which a decision maker cannot detect the option with the higher expected value without inspecting all attributes—he or she can only choose this option by chance (but also note there is a 50% baseline chance of choosing the EV–maximizing option in choice environments with two alternatives). The link between processing strategies and choice may help explain why participants in the uncorrelated condition chose the option offering the higher expected value more often.

Interim conclusion

We hypothesized risk–reward structures to systematically affect how people process decisions under risk. Our data suggest that people do not use the risk–reward structure to infer one attribute from the other (payoffs from probabilities or probabilities from payoffs) in decisions under risk, as they do in decisions under uncertainty (Pleskac and Hertwig, 2014; Leuker et al., 2018a; Skylark and Prabhu-Naik, 2018). Instead, learning about risk–reward structures affected decisions under risk more gradually, and depending on two aspects: Time available to participants and the type of risk–reward structure that was learned. Under timepressure (when they needed to respond within $1.5s$), participants in the negative risk–reward condition inspected slightly fewer attributes compared to participants in the other two conditions, but these differences were not credible. If anything, inspecting more attributes in the uncorrelated and positive conditions could imply an attempt to process the same information at a higher rate. Under no timepressure, when participants were asked to make the best possible choice, processing strategies changed gradually—namely such that participants in the negatively correlated risk–reward environment decided sooner among the options and inspected fewer attributes during this time. While participants in the negative condition shifted their attention to fewer attributes, participants in the uncorrelated condition inspected more attributes and divided their attention more evenly across them. This was also linked to participants in the negative risk–reward environment to choose the higher EV option less often than participants in the uncorrelated risk–reward environment.

These differences did not generalize to participants who had been exposed to a positive risk–reward environment. These participants maximized expected values less (Figure 3A), but also took more time than participants in the negative risk–reward environment. We originally hypothesized that participants may use less rigorous processing strategies after learning about *any* correlation between risks and rewards—however, processing strategies in the positive condition reflected more those of the uncorrelated condition. One intuition behind not observing frugal processing strategies after learning about positive risk–reward structures is that participants in the negative, but not positive, risk–reward environment could uphold their belief that risks and rewards are inversely related in the choice phase because they experienced a local risk–reward relationship (per trial) with fairly similar EVs. Participants in the positive condition, however, could have expected an environment with dominated options—if the positive risk–reward relationship had been maintained.

In sum, participants seemed to satisfice sooner in sampling information from attributes after they had been exposed to an environment in which risks and rewards were inversely related; and participants seemed to maximize more after they had been exposed to an environment in which risks and rewards were uncorrelated. Satisficing sooner seemed to come at the cost of choosing the higher EV option less often. To better synthesize these results, we used a computational modeling approach that captured how gaze and properties of the gamble together shaped preferences. In particular, the computational model revealed which aspects of the choice process were impacted by the risk–reward environment.



$$\text{where } \delta_{i,p,t,r} = \delta 0_p + \delta EV_{i,p,t,r} * (EV_{higher} - EV_{lower}) + \delta gaze_{i,p,t,r} * (gaze_{higher} - gaze_{lower})$$

Figure 4. Depiction of the Drift Diffusion Model and conditions for which parameters were estimated. We set the bias parameter β to .5 because participants cannot be biased to either the higher or lower EV option before inspecting it. The drift rate was a function of individual differences in ability/effort to detect the higher EV option (intercept), an EV coefficient δ_{EV} measuring how much (the use of) EV differences contributed to the drift rate and a gaze coefficient δ_{gaze} measuring how much gaze differences or biases contributes to choice.

Computational model of attention, response times and choice

We modeled our data using an extended drift diffusion model. Drift diffusion models have the advantage of making process predictions of how people form a preference for one options over the other in risky choice (Busemeyer and Townsend, 1993; Krajbich et al., 2010). According to these models, the construction of preference is an evidence accumulation process where over time, as people consider which gamble to choose, they accumulate evidence for or against each option (Figure 4). Once evidence reaches a threshold (accumulation models assume one threshold for each option), people choose accordingly. How far apart people set these thresholds determines how quickly they reach a decision: the closer they set the thresholds the shorter it will take to reach a decision, but the more variable people will be in the choices they make. Thus, the DDM can help explain how the formation of preferences is affected by, for instance, timepressure.

Researchers have identified two factors that determine how people accumulate information when making preferential choices: The value of the options (in gambles approximated by the gambles' EV), gaze or differences in gaze to the alternatives (Cavanagh et al., 2014; Stewart et al., 2016) and/or an interaction between the two (Krajbich et al., 2010, 2012; Krajbich and Rangel, 2011). To formally test for the influence of these factors, we fit an extended drift diffusion model in which the drift rate δ—that describes the average strength of evidence in each sample—was divided into two subcomponents, an EV coefficient δ_{EV} and a gaze coefficient δ_{gaze}. These two coefficients were modeled as free parameters and estimated given option characteristics (EV differences) and gaze patterns in a given trial. A parameter value that exceeds 0 (i.e. is positive) indicates that the coefficient contributes to the evidence accumulation process towards the higher EV option; a paramater value around 0 indicates little or no influence on the evidence accumulation process.

The model has two additional free parameters. One is the threshold separation α that denotes response caution. We predicted $alpha$ to be high in our "best" conditions and low in our "fast" conditions. The other one is the nondecision time τ that models the part of the response time that is unrelated to the pro-

Parameter	Description
Threshold separation (α)	Response caution/Threshold. Here, the top boundary is set to the higher EV option. In the best instruction, the two thresholds are set far apart producing slower decisions but a higher rate of expected value maximizing choices. In the fast instruction, boundaries are expected to be closer to the starting point producing faster decisions and more errant decisions that deviate from expected value maximizing choices.
Relative start point (β)	Response bias. If responses are unbiased, $\beta = 0.5$. Here, the higher EV option is not detectable from the beginning and the location of options has been counterbalanced. We therefore fix $\beta = .5$.
Nondecision time (τ)	Tau models the part of the response time that is unrelated to the processing of the option itself (e.g. encoding, motor response).
Drift rate (δ)	The average strength in evidence at each unit of time, with $\infty < \delta < \infty$. The sign of the drift rate indicates the average direction of the incoming evidence, with positive values indicating evidence in favor of the higher EV option. The magnitude of the drift rate characterizes the quality of the incoming information. Here, the drift rate is not a free parameter, but a linear function of the difference in expected values and the difference in the amount of time each gamble was fixated on (see below for coefficients).
EV Coefficient (δ_{EV})	Main effect of EV differences between higher and lower EV gambles on the drift rate.
Gaze Coefficient (δ_{gaze})	Main effect of gaze differences between higher and lower EV gambles on the drift rate.

Table 1. Parameters of the extended Drift Diffusion Model.

cessing of the option itself, e.g. motor encoding (no predictions as to the influence of different risk–reward environments). The parameters are summarized in Table 1. We obtained estimates of the parameters by fitting the extended DDM to the observed distributions of choices and response times using a Bayesian hierarchical implementation. Details and a comparison to other models, including a plain DDM without any drift-rate dependencies can be found in Supplementary Material (Table C3).[8]

Computational modeling results

Figure 5 displays the group level estimates of the nondecision time τ, threshold separation α and the two coefficients that determined the drift rate, the EV coefficient δ_{EV} and the gaze coefficient δ_{gaze}. Following the behavioral results, we focus on comparisons between the negative and uncorrelated risk–reward condition when discussing differences between parameters.

Threshold α

Consistent with response times, the threshold separation was credibly higher in the best ($M = 1.44$ $[1.27, 1.60]$) compared to the fast conditions ($M = 1.80$ $[1.20, 2.30]$). Under strong time pressure, boundaries are often hit by mistake and this leads to lower proportions of EV–maximizing choices—as observed in the behavioral data. Moreover, in the best condition, consistent with the behavioral data, participants in the negative condition set a lower threshold than participants in the uncorrelated condition ($M_{neg.>unc.} = -0.56$ $[-.86, -.24]$). This was not the case in the fast condition (Figure 5A). We also assessed how well the estimated parameters were in line with individual differences in participants' (descriptive)

[8] Briefly, we also included a model that consisted of additive *and* an interaction effect between gaze and value in our model comparisons, which had a worse fit in terms of DICs. We did not include an attentional DDM (aDDM)–like model in which the drift rate *solely* depended on the interaction between value and gaze without additive effects (as in Cavanagh et al., 2014)—the reason being that we wanted to partial out and compare "EV usage" for the three risk–reward conditions, which is not possible in an interaction–only model.

choices. To do so, we retrieved the individual parameter estimates (mean of the posterior distribution) and computed the proportion of trials in which participants chose the higher expected value option, and predicted choices from parameter estimates. This analysis revealed that individuals who set higher thresholds chose the higher expected value option more often, with a more pronounced link in the best condition ($M_{best} = 0.18$ $[0.08, 0.28]$, $M_{fast} = 0.07$ $[0.05, 0.09]$, plots are shown in Figure C3).

Nondecision time τ

As Figure 5B shows, nondecision times were higher in the best than in the fast conditions ($M_{best.>fast.} = 0.19$ $[0.07, 0.31]$), and there were no reliable differences between risk–reward conditions.

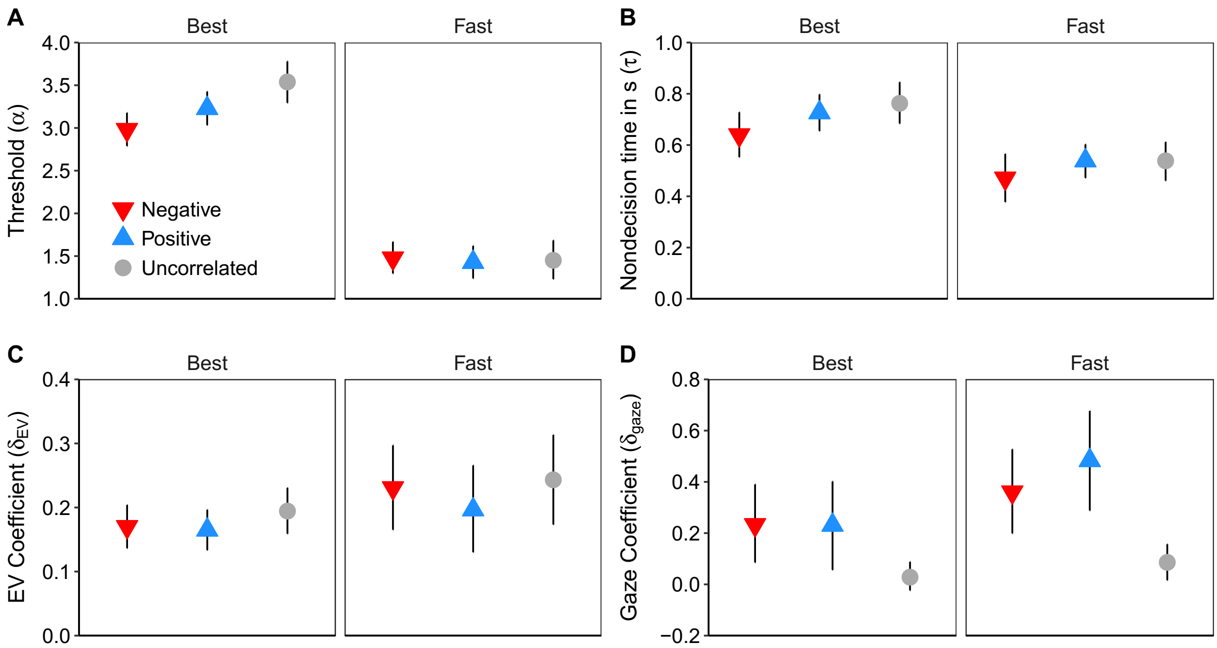


Figure 5. Parameter estimates after incidental learning. Group means and 95% Highest Density Intervals of the posterior distributions.

EV coefficient δ_{EV}

Figure 5C shows that EV differences between the higher and lower value gambles played a similar role under both timepressure conditions. Values consistently exceeding 0 indicate that EV differences reliably impacted the drift rate in all conditions. If anything, EV differences were slightly more important in the evidence accumulation process under timepressure—but also note the larger HDI, suggesting that the EV coefficient was more variable between individuals under timepressure. There were no differences in the EV coefficient across risk–reward environments. We again linked individual differences in parameter estimates to behavioral data, and found that a larger EV coefficient δ_{EV} was linked to a higher proportion of expected value–maximizing choices in the best instruction ($M_{best} = 1.06$ $[0.42, 1.71]$). There was a similar tendency in the fast condition, but the credible interval included 0 ($M_{fast} = 0.67$ $[-0.06, 1.40]$). This fits with the behavioral data in which participants did not inspect all of the attributes when processing demands were high—making it hard to detect expected value differences in this condition.

Gaze coefficient δ_{gaze}

As a comparison between panels in Figure 5D shows, the gaze coefficient or gaze bias was larger in the fast conditions, across risk–reward environments, but the difference between the different timepressure instructions conditions was not reliable. When comparing the gaze coefficient across risk–reward environments, there were some systematic differences. Specifically, for participants in the uncorrelated condition, gaze contributed little to the evidence accumulation process in both the best and the fast instructions—as indicated by a gaze coefficient close to (and including) 0 ($M\delta_{gaze} = 0.05$ $[-0.02, 0.14]$). By comparison, parameter estimates were above 0 in the negative and positive conditions, which indicate that how participants distributed their gaze across the options reliably impacted the drift rate in these conditions. That is, participants in the negative risk–reward environment relied more on gaze in their evidence accumulation than participants in the uncorrelated condition, across timepressure manipulations (best: $M_{neg.>unc.} = 0.20$ $[0.05, 0.37]$, fast: $M_{neg.>unc.} = 0.27$ $[0.10, 0.45]$). As Figure 5D shows, the parameter estimates in the positive condition are fairly similar to those in the negative condition.

We again linked differences in parameter estimates to individual differences in choice. Here, participants described by a larger gaze coefficient (δ_{gaze}) chose the higher expected value option less often ($M_{best} = -0.09$ $[-0.15, -0.03]$; $M_{fast} = -0.02$ $[-0.06, 0.03]$). The gaze coefficient was lowest in the uncorrelated condition, reflecting a lower impact of gaze differences on the drift rate—in other words, participants who distributed their attention more evenly (indicated by a gaze coefficient of 0) achieved the largest proportion of expected–value maximizing choices. Overall, these results are consistent with participants in the uncorrelated condition distributing their attention more evenly across payoffs vs. probabilities and inspecting more AOIs (Figure 3), thereby potentially achieving a higher proportion of expected–value maximizing choices.

Summary

To better characterize how participants accumulated evidence given previously learned risk–reward structures and timepressure, we fit an extended drift diffusion model in which the drift rate δ was divided into two subcomponents, an EV coefficient δ_{EV} and a gaze coefficient δ_{gaze}. In addition, we estimated condition–dependent variations in the threshold α and in the nondecision time τ. Model fits suggest that condition–dependent differences can be traced back to two aspects of the accumulation process. First, participants in the negative risk–reward condition satisficed sooner and set lower response thresholds (α), when making the best possible choice was emphasized. Second, choices in the correlated conditions were impacted more by how gaze was distributed among the options (δ_{gaze}), whereas this did not matter much for participants in the uncorrelated condition who—as the behavioral data suggested—may have distributed their attention more evenly across the options.

General Discussion

> If utilities and subjective probabilities are not independent, then there is no hope of predicting risky decisions unless their law of combination is known. (Edwards, 1954, p. 400)

What Edwards referred to was then–recent evidence that subjective probability weighting is influenced by the utilities in the bets (Irwin, 1955). He considered this possibility to be “most disturbing”. By now there is substantive evidence that the distributions of payoffs and probabilities in a given choice environment can influence peoples judgments and decisions (Lichtenstein and Slovic, 2006; Stewart et al., 2006; Quiggin, 1993; Ungemach et al., 2011; Parducci, 1965; Ludvig et al., 2014). This may not be as disturbing as Edwards predicted—but instead may challenge the way in which risky decisions should be studied. Here, our goal was to give a more complete account of how people make decisions under risk, by mimicking and exposing participants to different risk–reward structures, including an inverse relationship between risks and rewards as probably the most representative structure in nonlaboratory domains (Pleskac and Hertwig, 2014).

At the same time, risk–reward structures can be relatively uncorrelated in newly forming markets that have not yet reached an equilibrium. They have also been treated as uncorrelated in empirical studies of risky choice—which, due to the prominence of negative risk–reward structures—may only have captured a small fraction of how people make decisions under risk. It is possible that risk–reward structures do not impact choice at all. In this case, the same theories and conclusions—that have been by and large derived from decision making studies with independent, uncorrelated risks and rewards—apply to cases in which there is a systematic relationship between risks and rewards. However, theories of adaptive cognition make different predictions; namely, that the environment is represented or reflected in the mind, and these representations of the environment in the mind systematically affect how it operates (e.g., Anderson and Schooler, 1991; Brunswik, 1944; Gibson, 1979; Stewart et al., 2006; Ungemach et al., 2011; Shepard, 1987). In the case of correlated risks and rewards, people can decide to rely on a subset of cues in the environment because cues are interrelated (Brunswik, 1952). People, in fact, exploit risk–reward structures in decisions under uncertainty—in form of a risk–reward heuristic—to infer probabilities directly from the magnitude of the payoffs (Pleskac and Hertwig, 2014; Leuker et al., 2018a), or to infer payoffs from probabilities (Skylark and Prabhu-Naik, 2018). Based on these Brunswikian predictions, we investigated to what extent the relationship among risks and rewards impacts how people process options and form preferences in decisions under risk. We summarize our results and their boundary conditions, before we discuss their implications for risky decision making in the wild as well as their implications for laboratory studies of risky choice.

How learning about risk–reward structures affects decisions under risk

Just as people exploit risk–reward structures in decisions under uncertainty, namely by inferring probabilities directly from the magnitude of the payoffs (Pleskac and Hertwig, 2014; Leuker et al., 2018a), or inferring payoffs from probabilities (Skylark and Prabhu-Naik, 2018), people may use more frugal processing strategies if they can infer or expect the probability they will observe for a given payoff. Data from the current experiment partially supports this conjecture. While risk–reward structures did not lead people to ignore one of the attributes completely, participants who learned about a negative risk–reward relationship processed the options more frugally—notably when making the best possible choice was emphasized. Specifically, participants in the negative risk–reward environment shifted their attention more to payoffs compared to probabilities, and they responded faster—and less cautiously—than participants in the other two conditions. Participants in the uncorrelated environment distributed their attention most evenly. This was the case both within the gamble as seen from the proportion of gaze to payoffs and probabilities; and between gambles as evident in the gaze coefficient derived from the diffusion model. Moreover, they took more time to respond more cautiously than participants in the negative environment. There were some boundary conditions for these results. Specifically, learning about positive risk–reward environments did not lead to lower thresholds or inspecting fewer attributes—which may speak to positive risk–reward relationships being counter-intuitive given the many real world environments in which the highest rewards *are not* also the most likely ones. We found no robust evidence that people rely more on risk–reward structures when processing demands need to be reduced (also see Leuker et al., 2017, for a similar experiment with moderate timepressure, 2.5s).

Since choices were between two nondominated options, and if choosing the higher expected value option is taken as a normative approximation of rational choice, participants still paid a "simplification premium" for inspecting less information in the "best" environments. This is consistent with earlier simulations in which people can only maximize expected value choices and simultaneously rely on noncompensatory processing strategies in environments with dominated options (Payne et al., 1988). Conversely, participants in the uncorrelated condition, who relied on more rigorous processing strategies, maximized expected values more than participants in the negative condition. Following Simon (1978), people may have been well aware that searching less may compromise the likelihood with which they find and select the objectively better option and still satisfice. He referred to this as "procedural rationality" (p. 13). After all, attention and time are scarce resources, and people seem to satisfice sooner when they believe that environments allow them to do so, and the 'marginal benefits' of searching more are subjectively low.

We should again note that the experiment reported here followed an earlier eyetracking experiment with a similar design (Leuker et al., 2017). In this earlier work, we maintained a condition–dependent risk–reward environment across both the learning and the choice phase. Doing so confounds EV differences, with the risk–reward structure in the choice phase (with tiny EV differences in the positive environment and large ones in an uncorrelated environment). In the positive condition, maintaining risk–reward structures also means small tradeoffs within each gamble pair if nondominance is desired (e.g. 10 with $p = .1$ vs. 11 with $p = .09$; 80 with $p = .8$ vs. 75 with $p = .15$). Therefore, in the current experiment, the majority of gambles in the choice phase was identical across risk–reward conditions, and necessarily mismatched the previously learned risk–reward environment (e.g. the common gambles included subsets of gamble pairs with high payoffs/high probabilities; low payoffs/low probabilites—both of these are unexpected in an environment in which risks and rewards are otherwise inversely related). We expected participants to vary their processing strategies depending on whether these gambles matched or mismatched learned risk–reward structures (see H5 in preregistration), but because processing strategies in the choice phase were stable across different gamble types; we collapsed across them in our analyses. Thus, our conclusions from the experiment presented here pertain to the *priors* people have about a risk–reward environment (i.e. the risk–reward environment they assume or have been exposed to previously) rather than the risk–reward environment they are actually in. However, differences in processing strategies presented here are consistent with results presented in our earlier study, in which an inverse risk–reward relationship also resulted in fewer inspected attributes than an uncorrelated one. Taken together these results imply that the extent to which people satisfice or maximize by carefully comparing expected values, and ultimately select the normatively better option, is driven by the expectations they hold about a given risk–reward environment.

Is incidental learning special?

We also assessed to what degree our results generalize to another learning condition, in which N = 92 participants who learned about one of three risk–reward structures explicitly. The learning phase consisted of a function learning task in which participants "guesstimated" the probabilities linked to a range of different payoffs until they reached a learning criterion. Since all participants in the correlated conditions reached this criterion, all of them were aware of the condition–specific risk–reward relationship. Participants learned relatively quickly in the correlated conditions (for details see Supplementary Material), whereas most participants in the uncorrelated condition completed all 100 trials because there was no function to learn. Briefly, not all of the results from the incidental learning condition generalized to explicit learning: After explicit learning did not lead participants in the negative risk–reward environment to lower their decision thresholds (α), evidence accumulation was driven by gaze in all three risk–reward environments (δ_{gaze}). These results are reported in detail in the Supplementary Material.

What is the conceptual difference between explicit and incidental learning? One way to understand these differences is to think back to what participants could learn in incidental learning conditions when pricing gambles from different risk–reward environments. They learn two elements of risk–reward environments: First, they can learn about the functional form of the risk–reward structure (are risks and rewards positively related, negatively related or uncorrelated?). Second, they can also learn about the expected value distributions in a given environment, with a negative risk–reward structure being composed of many options with similar values ($20 with p = .8, and $80 with p = .2). This resembles the structure of the environment outside the lab, in which EVs across options are typically identical in monetary domains with a pay–to–play structure. For instance, the probability of obtaining a reward when there is a $1 pay–to–play fee is given by $p = 1/(1 + gain)$. This mechanism is, for example, found in roulette, and gives rise to risks and rewards being inversely related through a power law. In the current experiments, we exposed people to a linear relationship between risks and rewards and thus intermediate values ($50E\$$) had slightly higher EVs than high and low values.

While the functional form can be learned equally well—if not better—in an explicit learning task, explicit learning requires an additional step to then infer what that means for EV distributions in a given choice environment. It may be the perception of similar EV distributions that led participants in the negative implicit condition to set lower thresholds and sample less information. This is consistent with other work which indicates that negative risk–reward environments can elicit "EV surprise" for oddballs with higher expected values than usually experienced (Leuker et al., 2018b). In sum, explicit learning seems to be systematically different from incidental learning, potentially because incidental learning conveys a different type of information about the environment (namely a distribution of possible expected values *and* the risk–reward relationship, i.e. is it positive or negative?) than explicit learning (primarily the risk–reward relationship).

Implications for theories of risky choice

In understanding decision–making under risk, these findings provide a new perspective on why people may select more or less effortful choice strategies (Payne et al., 1988). It is sometimes assumed that a decision–maker has a repertoire of well–defined strategies that he chooses among by considering the expected costs and benefits of each strategy (Rieskamp and Otto, 2006; Payne et al., 1988). The selection process could be a conscious process of applying a meta-strategy or an unconscious decision triggered by experience (Payne et al., 1993). Here, we showed that more or less effortful processing indeed depends on the expectations a decision maker holds about her environment, and the experiences made in it. If a risk–reward environment is assumed to be uncorrelated, participants may exert more effort than in an environment for which they assume an inverse relationship among risks and rewards. In many choice environments outside the lab, risks and rewards are inversely related. In these environments, there may not be a normatively better (higher EV) option, and the decision maker needs to decide between similar options on other grounds (e.g. risk preferences, stakes, or even choose randomly).

The results also showed that probability information is not ignored if it is, in principle, available. It may be an adaptive strategy to “keep an eye” on probability information, for instance to detect above–average options in risky choice in newly forming markets—in which a risk–reward structure is not (yet) in place (Pleskac and Hertwig, 2014; Pleskac et al., prep).

Implications for testing risky choice theories in the lab

What do our results mean for researchers who wish to understand risky decision making? As Figure 1 shows, many key results are derived from laboratory studies with globally uncorrelated risks and rewards. Here, we tested how *learning* about different risk–reward structures affects subsequent decisions under risk with a majority of gambles being common across conditions (and thus uncorrelated). Modeling these choices with Prospect Theory (see Supplementary Figure C9) suggests that the different risk–reward structures did not systematically influence participants subjective preferences—as indicated by similar value function and probability weighting parameters[9]. Consistent with results of the extended drift diffusion model, the parameter estimates showed that participants in the negative risk–reward environment made less deterministic choice compared to the other two risk–reward conditions.

One limitation of this experiment is the large amount of common gamble pairs used, which also means that participants can adjust their expectations of a previously learned risk–reward structure throughout the experiment (for instance back to assuming that risks and rewards are globally uncorrelated), whereas condition–dependent gamble pairs may help to reinforce the previously learned structure. In this case, one way to test the influence of risk–reward structures in decisions under risk would be to intersperse only a few critical gamble pairs into an environment with negatively, positively or uncorrelated gamble pairs

[9]Not reported in the main manuscript because Prospect Theory does not make specific predictions about the evidence accumulation process and how it impacts choices and response times—as diffusion models do (Ratcliff and McKoon, 2008), for specifications and parameter estimates see Supplementary Material C9)

(Birnbaum and Chavez, 1997). We used such an approach in an earlier study by adding gambles that are known to produce the certainty effect (Kahneman and Tversky, 1979) to environment gambles with different risk–reward structures (Leuker et al., 2018a)—expecting that risk–reward environment may lower attention to probabilities (100% vs. 80%) and reducing the proportion of participants who will opt for the sure option. However, the certainty effect was robust across risk–reward environments. This is consistent with results from the current experiment suggesting that probability information is never fully ignored as a function of risk–reward structures. Thus, the key conclusions drawn from critical gambles in decisions under risk seem are replicable and therefore seem to be internally valid. However, a choice between $4000 with a probability of 80% or $3000 for sure still seems "more like a mere homunculus of the laboratory out in the blank" (1955, p. 204)—potentially even triggering more attention than people often spend in the wild where these payoffs are usually unlikely. More generally, it has been found that people are extremely risk–averse for such high payoffs (Holt and Laury, 2002), potentially reflecting an unwillingness to gamble in opportunities they seldom experience.

Conclusion

In 1953, Maurice Allais showed that people prefer a 100% chance of winning 100 million Fr. over composite alternative that offered a 10% chance of winning 500 million Fr., a 89% chance of 100 million Fr. and a 1% chance of winning nothing at all. Most of us do not get to make such choices very often. This poses the question of whether theories of risky decision making have overrelied on gambles that are like a mere homunculus of the laboratory out in the blank (Brunswik, 1955), and failed to characterize risky choices in more representative risk–reward environments. On the one hand, choices and subjective preferences (i.e. how people weigh payoffs and probabilities) are not impacted by priors about the global risk–reward structure—rendering the preferences and theories elicited in the what Savage called "small worlds" (1954) of risky choice tasks *internally* valid. On the other hand, the risk–reward environment people assume to be in affects how they process options in risky choice. Specifically, exposure to gambles in which risks and rewards are uncorrelated produces leads to more cautious responding compared to exposure to an environment where risks and rewards were inversely related. This suggests that an environment where risks and rewards are uncorrelated, such as in newly forming markets and laboratory risky choice studies, may lead people to maximize while environments in which risks and rewards are inversely related permit, and appear to promote, satisficing.

References

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'ecole americaine. *Econometrica*, 21(4):503–546.

Anderson, J. R. and Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 2(6):396–408.

Becker, G. M., Degroot, M. H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, 9(3):226–232.

Bernoulli, D. (1738/1954). Exposition of a new theory on the measurement of risk. *Econometrika*, 22:23–36.

Birnbaum, M. H. (2011). Testing Theories of Risky Decision Making Via Critical Tests. *Frontiers in Psychology*, 2:315.

Birnbaum, M. H. and Chavez, A. (1997). Tests of theories of decision making: Violations of branch independence and distribution independence. *Organizational Behavior and Human Decision Processes*, 71(2):161–194.

Boole, G. (1854). *Investigation of the Laws of Thought*. Macmillan and Co., Cambridge, UK.

Brunswik, E. (1944). Distal focussing of perception: Size constancy in a representative sample of situations. *Psychological Monographs*, 56(254):1–49.

Brunswik, E. (1952). The conceptual framework of psychology. *Journal of Consulting Psychology*, 16(6):475–475.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62(3):193–217.

Busemeyer, J. and Townsend, J. T. (1993). Decision Field Theory: A dynamic cognitive approach to decision making. *Psychological Review*, 100:432–459.

Cavanagh, J. F., Wiecki, T. V., Kochar, A., and Frank, M. J. (2014). Eye Tracking and Pupillometry are Indicators of Dissociable Latent Decision Processes. *Journal of Experimental Psychology: General*, 143(4):1476–1488.

Dalmaijer, E. S., Mathôt, S., and Van der Stigchel, S. (2013). PyGaze: An open-source, cross-platform toolbox for minimal-effort programming of eyetracking experiments. *Behavior Research Methods*, pages 1–16.

Daston, L. (1988). *Classical probability inthe enlightement*. Princeton University Press, Princeton, NJ.

David, F. N. (1962). *Games, Gods, and Gambling: A history of pf probability and statistical ideas*. Dover Publication, Mineola, NY.

Dhami, M. K., Hertwig, R., and Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. 130(6):959–988.

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin*, 51(4):380–417.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics*, 75:643–669.

Fiedler, S. and Gloeckner, A. (2012). The dynamics of decision making in risky choice: An eye-tracking analysis. *Frontiers in Psychology*, 3(OCT):1–18.

Fox, C. R. and Clemen, R. T. (2005). Subjective probability assessment in decision analysis: Partition dependence and bias toward the ignorance prior. *Management Science*, 51(9):1417–1432.

Fox, C. R. and Rottenstreich, Y. (2003). Partition priming in judgment under uncertainty. *Psychological Science*, 14(3):195–200.

Gelman, A. and Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7(4):457–511.

Gibson, J. (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston, MA.

Gigerenzer, G., Hoffrage, U., and Kleinbolting, H. (1991). Probabilistic Mental Models: A Brunswikian Theory of Confidence. *Psychological Review*, 98(4):506–528.

Gloeckner, A. and Herbold, A.-K. (2011). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making*, 24(1):71–98.

Hacking, I. (1975). *Emergence of Probability*. Cambridge University Press, Bristol.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5):1644–1655.

Irwin, F. W. (1955). Stated Expectations as Functions of Probability and Desirability of Outcomes. *Journal of Personality*, 21(3):329–335.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Krajbich, I., Armel, C., and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, 13(10):1292–1298.

Krajbich, I., Lu, D., Camerer, C., and Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in Psychology*, 3.

Krajbich, I. and Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, 108(33):13852–13857.

Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A tutorial with R, JAGS, and Stan, 2nd Ed.* Academic Press/Elsevier, New York, NY.

Leuker, C., Pachur, T., Hertwig, R., and Pleskac, T. J. (2018a). Exploiting risk–reward structures in decision making under uncertainty. *Cognition*, pages 186–200.

Leuker, C., Pachur, T., Hertwig, R., and Pleskac, T. J. (2018b). Too good to be true? Psychological responses to surprising options in risk–reward environments (Manuscript submitted for publication).

Leuker, C., Pleskac, T. J., Pachur, T., and Hertwig, R. (2017). How the mind exploits risk-reward structures in decisions under risk. In Gunzelmann, G., Howes, A., Tenbrink, T., and Davelaar, E., editors, *Proceedings of the 39th Annual Conference Cognitive Science Society*, pages 2543–2548, London, UK.

Lichtenstein, S. and Slovic, P. (2006). *The Construction of Preferences*. Cambridge University Press, New York, NY.

Lopes, L. L. and Oden, G. C. (1999). The Role of Aspiration Level in Risky Choice: A Comparison of Cumulative Prospect Theory and SP/A Theory. *Journal of Mathematical Psychology*, 43(2):286–313.

Ludvig, E. A., Madan, C. R., and Spetch, M. L. (2014). Extreme Outcomes Sway Risky Decisions from

Experience. *Journal of Behavioral Decision Making*, 27(2):146–156.

MacCrimmon, K. R. and Larsson, S. (1979). *Utility Theory: Axioms Versus 'Paradoxes'*, pages 333–409. D. Reidel Publishing Company.

Manohar, S. G. and Husain, M. (2013). Attention as foraging for information and value. *Frontiers in Human Neuroscience*, 7(11):711.

Marr, D. (1982). Vision: A computational approach.

Mellers, B. A., Schwartz, A., Ho, K., and Ritov, I. (1997). Decision Affect Theory: Emotional Reactions to the Outcomes of Risky Options. *Psychological Science*, 8(6):423–429.

Pachur, T., Hertwig, R., Gigerenzer, G., and Brandstätter, E. (2013). Testing process predictions of models of risky choice: A quantitative model comparison approach. *Frontiers in Psychology*, 4.

Pachur, T., Schulte-Mecklenbeck, M., Murphy, R. O., and Hertwig, R. (2018). Prospect theory reflects selective allocation of attention. *Journal of Experimental Psychology: General*, 147(2):147–169.

Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review*, 72(6):407–418.

Payne, J. W., Bettman, J. R., and Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3):534–552.

Payne, J. W., Bettman, J. R., and Johnson, E. J. (1993). *The Adaptive Decision Maker*. Cambridge University Press, New York, NY.

Pleskac, T. J., Conradt, L., Leuker, C., and Hertwig, R. (in prep). The ecology of competition: A linchpin for risk–reward structures. *Manuscript in prepartion.*

Pleskac, T. J., Diederich, A., and Wallsten, T. S. (2015). Models of decision making under risk and uncertainty. In Busemeyer, J. R., Townsent, J. T., Wang, Z. J., and Ami, E., editors, *The oxford handbook of computational and mathematical psychology*, pages 209–231. Oxford Library of Psychology.

Pleskac, T. J. and Hertwig, R. (2014). Ecologically rational choice and the structure of the environment. *Journal of Experimental Psychology: General*, 143(5):2000–2019.

Quiggin, J. (1993). Rank-Dependent Expected Utility: An Outline. In *Generalized Expected Utility Theory: The Rank-Dependent Model*, pages 55–74.

Ratcliff, R. and McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, 20(4):873–922.

Rieskamp, J. and Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, 135(2):207–236.

Savage, L. J. (1954). *The Foundations of Statistics*. John Wiley Sons, New York, NY.

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, 237(4820):1317–23.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63(2):129–138.

Simon, H. A. (1978). Rationality as process and as product of thought. *The American Economic Review*, 68(2):1–16.

Skylark, W. J. and Prabhu-Naik, S. (2018). A new test of the risk-reward heuristic. *Judgment and Decision Making*, 13(1):73.

Stan Development Team (2016). RStanArm: Bayesian applied regression modeling via Stan (version 2.9.0-4).

Stewart, N., Chater, N., and Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, 53(1):1–26.

Stewart, N., Hermens, F., and Matthews, W. J. (2016). Eye movements in risky choice. *Journal of Behavioral Decision Making*, 29(2-3):116–136.

Thorngate, W. (1980). Efficient decision heuristics. *Behavioral Science*, 25(3):219–225.

Todhunter, I. (1865). *A History of the Mathematical Theory of Probability from the Time of Pascal to that of Laplace*. Macmillan and Co., Cambridge and London.

Ungemach, C., Stewart, N., and Reimers, S. (2011). How incidental values from the environment affect decisions about money, risk, and delay. *Psychological Science*, 22(2):253–260.

Von der Malsburg, Titus (2015). Saccades version 0.1-1. cran.r-project.org/web/packages/saccades/.

von Neumann, J. and Morgenstern, O. (1947). *Theory of Games and Economic Behavior*. Princeton University Press (1980), Princeton, NJ.

Zur, H. B. and Breznitz, S. J. (1981). The effect of time pressure on risky choice behavior. *Acta Psychologica*, 47(2):89–104.

5 | When money talks: Judging risk and coercion in high-paying clinical trials

Leuker, C., Samartzidis, L., Hertwig, R. & Pleskac, T.J. *In Prep.*

Abstract

Millions of volunteers take part in clinical trials every year. This is unsurprising, given that clinical trials are often much more lucrative than other types of unskilled work. However, offering very high pay is sometimes considered ethically inappropriate. To investigate why, we asked 1,372 respondents to evaluate a hypothetical medical trial for a new Ebola vaccine offering three different payment amounts. Some individuals used very high pay as a cue to infer the potential risks a clinical trial poses. At the same time, despite the risk they may pose, some clinical trials may be too lucrative to be turned down. Both perceived risk and coercion shape how people evaluate clinical trials; they also help explain why there is a distaste for markets in which parties are remunerated for the risk they take—that is, why markets in which health and wellbeing are turned into commodities are considered repugnant.

This chapter is available as a preprint under a CC-BY 4.0 license | Preprint: osf.io/9wmta | Preregistration wave 1: osf.io/yftvh | Preregistration wave 2: osf.io/kumge | ACKNOWLEDGMENTS: We would like to thank Sandro Ambuehl, Thomas Wallsten, and Alvin E. Roth for comments on an earlier draft and Kate Pleskac for editing the survey questions.

Introduction

"All of us, with the exception of the independently wealthy and the unemployed, take money for the use of our body."

Martha Nussbaum, 1998

Would you volunteer to take part in a drug study that offered €1,900, travel expenses, and a two-week stay at a pharmaceutical research institute in return for swallowing a drug on 10 consecutive days, undergoing extensive medical tests, and providing at least 40 blood samples? In 2015, 128 people volunteered to do just that at a Biotrial research institute in France. For one volunteer, who died, and six others, who were hospitalized, the money was not worth it. Later, experts underlined "the astonishing and unprecedented nature" of these incidents, which were the result of a reaction in the brain "unlike anything seen before" (AFP, 2016). No such effects were observed in pretests with animals, despite doses 400 times stronger than those the human volunteers received (Randerson, 2016). If the researchers could not have anticipated the risks, how could the volunteers have suspected what they were signing up for? And even if volunteers had somehow intuited the risks, could the study have been too lucrative to be turned down?

These are important questions to ask given the role clinical trials play in the development of new pharmaceutical treatments. The transition from animal testing to phase 1 trials, which assess the safety and effectiveness of a particular substance in humans, is critical. Millions of volunteers, both healthy and with existing health conditions, are sought for clinical trials every year. While volunteers with existing conditions may take part in medical research studies in the hopes of improving their health, healthy volunteers find their motivation elsewhere: More than 90% say money is their main incentive for participating (Bigorra and Baños, 1990; van Gelderen et al., 1993). Clinical trial volunteers sign an informed consent that outlines the details of the study including potential risks, after which the research institute and the person offering their body engage in a voluntary market transaction. This transaction is, to some, no different from any other market transaction (Wilkinson and Moore, 1997).

To others, however, such a market transaction is hardly comparable to getting a haircut. In fact, clinical trial markets are sometimes placed in the same category as paid kidney donations and prostitution—*repugnant* transactions that people deem morally repulsive and therefore want to prevent (Roth, 2007). The repugnance of clinical trial markets has been attributed to coercive remuneration for volunteers, since payment in clinical trials is usually much higher than unskilled work, which often pay minimum wages. In 2015, minimum wage workers in France would have earned €729 in two weeks, less than 40% of what the volunteers earned (Fric, 2016).

In support of this, Ambuehl et al. (2015) found that some people, who we call "Doubters," rated a clinical trial offering $10,000 as *more* coercive than the same trial offering $1,000.[1] Consistent with this evaluation, Doubters also thought participants would regret enrolling in the higher paying study and be

[1] We call the "ethicists" of the original survey Doubters and refer to "economists" as Trusters to reflect the fact that Doubters are consistently more critical than Trusters of high-paying clinical trials.

better off not participating. Finally, the Doubters stated they would be less likely to approve such a study if they were on an institutional review board (IRB) panel. A second group of respondents, who we call “Trusters”, rated the trial offering \$10,000 as *less* coercive. They also thought that individuals would be better off if they took part and would not regret participating. Trusters on an imagined IRB panel were more likely to approve the study offering \$10,000 than the one offering \$1,000.

The sense of possible coercion may not be the only factor that discerns between these two groups. The key difference between clinical trial participation and other types of jobs is that clinical trials expose participants to unknown risks (McNeill, 1997). Volunteers may be compensated *after* side effects are experienced, but policies vary.[2] In other cases, the compensation volunteers receive for participating in the clinical trial may be considered to directly offset the risks to which they are exposed: Around one third of surveyed research institutes reported that one rule of thumb for determining pay is the anticipated risk participants incur (Dickert et al., 2002). This may explain why payment is sometimes treated as a cue indicating the risk of negative consequences from participating in a clinical trial (Cryder et al., 2010, though results are inconsistent, see Mantzari et al., 2014).

Consequently, people who judge the ethicality—and by extension, the repugnance—of a high-paying clinical trial may not only heed the coercion caused by the high payoffs, but also the risks that the payoffs signals. Very high payoffs offered in clinical trials could stress the potential harm to participants and thereby decrease people’s approval of the study. If so, the difference between the two groups of respondents in Ambuehl et al. (2015) may at least be partly due to different inferences about the implied risks. This is not an uncontested hypothesis. Ambuehl et al. (2015) maintained that the risks of the \$1,000 and \$10,000 payment schemes were seen as equal because respondents got the same description of the trial and saw all the possible payoff schemes (see their Footnote 7). Moreover, others have argued that high payment amounts may conceal the risk involved in taking part in a study (Bentley and Thacker, 2004; McNeill, 1997), or even impair prospective participants’ ability to think carefully about the risks and benefits involved (Casarett et al., 2002).

To investigate whether the ethicality of high-paying clinical trials is influenced by the risks inferred from the payoff magnitude, we presented online survey respondents with a hypothetical medical trial that compensated volunteers with 50, 1,000, or 10,000 (materials adapted from Ambuehl et al., 2015). Respondents estimated how many prospective participants would experience side effects, and evaluated the clinical trial on several other dimensions pertaining to coercion and ethicality. We asked four research questions: (1) Do the results from Ambuehl et al. (2015) replicate? (2) Do people perceive higher payment to be associated with higher risk? (3) Why do people consider high pay in clinical trials to be ethically inappropriate? and (4) How do payment-dependent inferences shape judgments on the repugnance of clinical trials?

[2] Member states of the European Union must offer systematic compensation for research-related injuries; no such regulation exists in the United States (Pike, 2012). Approximately 59% of institutions conducting clinical research in the United States explicitly offer conditional or unconditional compensation for research-related injuries (Resnik et al., 2014). Yet, even where laws exist, volunteers are often not fully aware of their financial vulnerability when they enroll in clinical trials (Manning, 2017).

Method

Participants

In total, N = 1,565 respondents completed our survey posted on Prolific Academic for a flat payment of 2.10. Inclusion criteria were fluency in English (self–assessed), and a minimum approval rate of 80% in earlier studies completed on the platform. The data were collected in two waves ($N = 354$ in wave 1, $N = 1,211$ in wave 2). The first sample size was determined based on the availability, and a reasonable allocation of funds (osf.io/yftvh), since our survey added several sets of questions, which lengthened the original 12-minute survey to an estimated 25 minutes. Wave 1 was smaller than the sample size in the original study ($N = 1,445$, Ambuehl et al., 2015). After examining the results, we sought additional evidence to test our predictions regarding the estimated side effects *between* respondents. We aimed to recruit an additional $N = 1,200$ respondents in wave 2 (osf.io/kumge). We improved our survey in wave 2 by asking respondents to assess how repugnant they find clinical trials in general. The total sample size of waves 1 and 2 roughly matched that of the original study. Inclusion criteria, predictions, survey questions and the reasoning behind collecting additional responses were preregistered. The surveys were approved by the IRB of the Max Planck Institute for Human Development. We collapsed our data in the main manuscript and report effects of wave 1 vs. wave 2 in the analyses of interest in the Supplementary Materials (in short: They were largely independent of wave). We analyzed data from N = 1,428 (N = 316 in wave 1, N = 1,112 in wave 2) respondents who passed three simple attention checks (questions pertaining to the instructions and the clinical trial description that could be found on the same page). The final sample consisted of 885 females, 535 males, and eight who identified as 'other,' and was on average 36.1 years old (range 18–78 years, SD = 11.8).

Survey

Vignette

We used a vignette from a study about a hypothetical clinical trial testing an Ebola vaccination (Ambuehl et al., 2015). Respondents were put in the position of an IRB panelist evaluating the trial, which was described as a phase 1 trial which sought to test the vaccination for the first time on 100 female volunteers, after pretests in rats and chimps. The vaccine was described as having "low, but nonzero risks" (see Supplementary Materials for full vignette). The clinical trial remunerated prospective participants with one of three payment amounts [50/1,000/10,000]. Each respondent saw the vignette with all three payment amounts, but the primary analyses in the manuscript are based on the first payment amount respondents saw (i.e., between-respondents). Within-respondent analyses were consistent with the between-respondent analyses and are reported in the Supplementary Materials. The original survey included U.S. residents recruited via Amazon Mechanical Turk and hence used U.S. currency ($). As our sample was British, we used British currency () and did not transform the values according to exchange rates.

Side effects

After reading the scenario, respondents were asked to assess how many of the 100 participants they expected to experience [any/mild/severe] side effects.

Clinical trial evaluations

As in the original survey, each respondent evaluated the clinical trial answering the following questions in order. (1) *IRB approval.* Suppose you are a member of the ethics committee that has to approve the institute's study with [*payment*]. How would you decide? (2) *Personal approval.* How much do you personally approve of the institute's proposal to enlist and compensate study participants from both rich and poor neighborhoods in this way? (3) *P (better off without).* [Description of A.S., a woman earning $1,500/month] Suppose that 10 women similar to A.S. see the institute's study participation invitation. How many of the 10 would be better off if the institute had never posted the study participation invitation?[3] (4) *P (enroll).* How many of the 10 do you think will eventually participate in the study in exchange for [*payment*]? (5) *Voluntariness.* If A.S. decides to participate in the study for [*payment*], how would you describe her decision? (Likert scale with extremes labelled "She was coerced" and "Her decision was entirely voluntary"). (6) *P (regret accepting).* If A.S. decides to participate in the study, how likely is it that she will later regret her decision? (7) *P (regret rejecting).* If A.S. decides NOT to participate in the study, how likely is it that she will later regret her decision?

Since *IRB approval* and *personal approval* were highly correlated, we focus on *IRB approval* in the main manuscript. Similarly, since *P (regret accept)* and *P (regret reject)* were highly correlated, we focus on *P (regret accept)* in the main manuscript (corresponding models for the other variables can be found in the Supplementary Materials).

Different types of respondents

At the end of the survey, participants rated the appropriateness of different payments side by side, on a 7-point Likert scale ("For each of the following ways of compensating study participants, please indicate how ethically appropriate you think it is. Recall that the study tests for effects of a vaccine, and although nobody expects such side effects to occur, if this were known, there would be no need to run a study. There is no special compensation if side effects occur."). This question, as in Ambuehl et al. (2015), was later used to categorize respondents into three types (Doubters, Trusters, and Others): Doubters' strictly preferred offering 1,000 over 10,000; Trusters' strictly preferred offering 10,000 to 1,000, and Others were indifferent to these two payment amounts. These names reflect that Doubters are more critical of clinical trials (in particular those that offer high pay) throughout, whereas the responses of Trusters are consistent with payoff-independent clinical trial *risk inferences.*

[3] In the original survey, respondents rated the ethicality of the trial for two women: one earning 1,500/month and one with a minimum-wage job. In our survey, participants only evaluated the trial for a woman earning 1,500/month. Since the minimum wage in the United Kingdom is around 1,250/month (Staff Writers, 2018), we did not expect results to differ substantially in these two conditions.

Repugnance

In our second wave of data collection, we also asked respondents to rate the "repugnance" of clinical trials in general, based on four items: "Clinical trial markets... (1) are deplorable, (2) morally permissible (R), (3) should be banned (R), (4) should be monitored (R)". Participants responded to these questions on a Likert scale from strongly agree (1) to strongly disagree (7). We recoded the answers to questions 2–4 (R) such that (1) denoted 'less repugnant' and (7) denoted 'more repugnant'. Based on these items, we computed a repugnance measure for each respondent (their average rating across the four items). Individual responses to questions 1–3 were highly correlated. Responses to question 4 ("should be monitored") were uncorrelated to the other responses. Since including and excluding question 4 from our repugnance measure led to qualitatively identical results, and we had not predicted it to be uncorrelated to the other measures a priori, we included it in the reported regressions. All "repugnance" analyses are based on wave 2 data—the repugnance questions had not been included in wave 1 (also see OSF).

Numeracy, risk-taking and demographics

Lastly, respondents completed the Berlin Adaptive Numeracy Test (Cokely et al., 2012), an 'insurance task' (not reported, but see OSF for wording) and items on their self-rated willingness to take risks. Six domain-specific risk-taking items were included (driving, financial, recreational, occupational, health, and social). The wording was as follows: "People can behave differently in different situations. How would you rate your willingness to take risks in (e.g.) *the financial domain*?". A general risk-propensity item was included in wave 2 ("How willing are you to take risks in general?", Dohmen et al., 2011). Respondents also indicated whether they had ever thought about participating in medical research themselves. The survey concluded with demographic questions (gender, age, education). The final sample, on average, took 22.5 minutes (range $8.7 - 74.5$ minutes, $SD = 10.38$) to complete the survey.

Statistical analyses

We relied on Bayesian estimation techniques (Kruschke, 2014) and applied Bayesian Generalized Linear Models using Stan in R for regression analyses with the rstanarm package (Stan Development Team, 2016). We ran three chains using Markov Chain Monte Carlo sampler to draw from posterior distributions of parameters, with 10,000–30,000 samples per chain (to ensure an effective sample size of >10,000 for each regressor), and a burn-in of 500 samples. We investigated the convergence of our posteriors through visual inspection and the Gelman-Rubin statistic (Gelman and Rubin, 1992). In general, we report the mean of the posterior distribution of the parameter or statistic of interest and two-sided 95% equal tail credible intervals (CI) around each value. In the replication section, our data analysis corresponds to the analysis in the original study.

Results

The results section is organized as follows: First, we present a qualitative replication of Ambuehl et al. (2015), again showing that Doubters and Trusters disagree about offering 10,000 when evaluating various

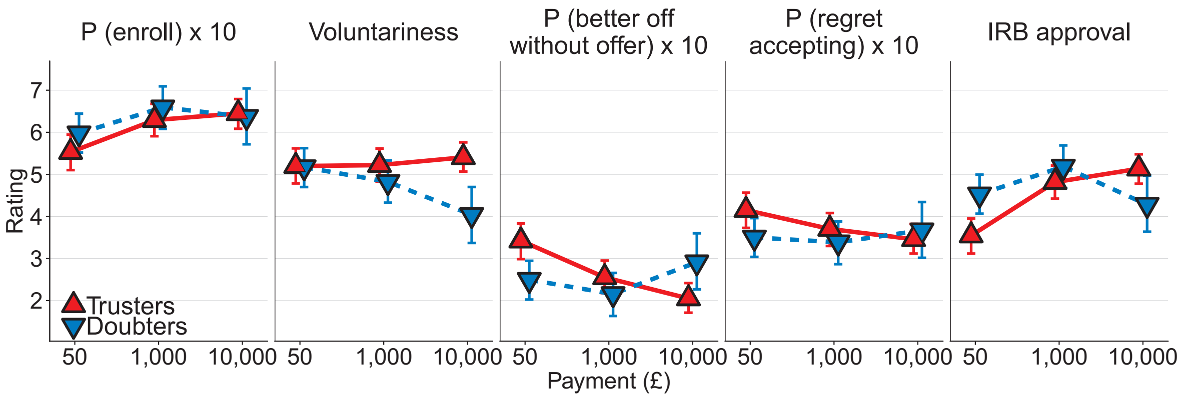


Figure 1. Responses of Doubters and Trusters for various payment amounts (between-respondents). Black triangles represent sample means. Colored triangles and error bars represent the means and the 95% highest density intervals of the posterior predictive distributions.

clinical trial dimensions. Second, we show how payment can leak information about the riskiness of clinical trials. Third, we investigate the extent to which risks and coercion were linked to clinical trial evaluations (i.e., linking sections 1 and 2). Lastly, we explore risk and coercion as determinants of repugnance for clinical trials more generally.

Clinical trial evaluations (replication of Ambuehl et al., 2015)

Following the original survey, we categorized respondents into three types, based on how they rated the ethicality of different payment amounts ($\delta_{rating} = rating_{10,000} - rating_{1,000}$). Trusters rated a payment of 10,000 as strictly more ethically appropriate than a payment of 1,000 ($N = 712$, 50%, $\delta_{rating} = 1.84$, $CI = [1.77, 1.91]$). Doubters rated a payment of 1,000 as strictly more ethical than a payment of 10,000 ($N = 378$, 27%, $\delta_{rating} = -2.24$, $CI = [-2.33, -2.14]$). A subset of respondents rated both payment amounts equally (Others, $N = 338$, 23%, $\delta_{rating} = 0.00$). The proportions of each type were highly similar to the proportions reported in Ambuehl et al. (2015). In the main results section we report analyses using responses to the first payment amount a respondent saw; each respondent therefore appears only once in the analyses (results for Others and within-respondent analyses shown in Supplementary Materials).

As in the original survey (Ambuehl et al., 2015), both groups (Trusters, and Doubters) evaluated clinical trials similarly as pay increased from 50 to $1,000$ (Figure 1): For instance, both Doubters and Trusters were more likely to give IRB approval for a trial offering 1,000 compared one offering 50, likely because the trial was described as requiring 40 hours of commitment, for which 1,000 seems fairer than 50. Figure 1 also shows that Trusters and Doubters expected approximately the same number of people to enroll in a trial that offered 10,000 compared to one that offered 1,000 (all CIs across payment amounts and interactions between groups included 0). This is different from the original survey, where both groups expected more people to enroll for 10,000.

However, as in the original survey, offering 10,000 rather than 1,000 to clinical trial volunteers was evaluated differently by Doubters and Trusters with regard to 'voluntariness,' 'being better off without the offer,' 'regret accepting' and—as a possible combination of the aforementioned concerns—'IRB approval.'

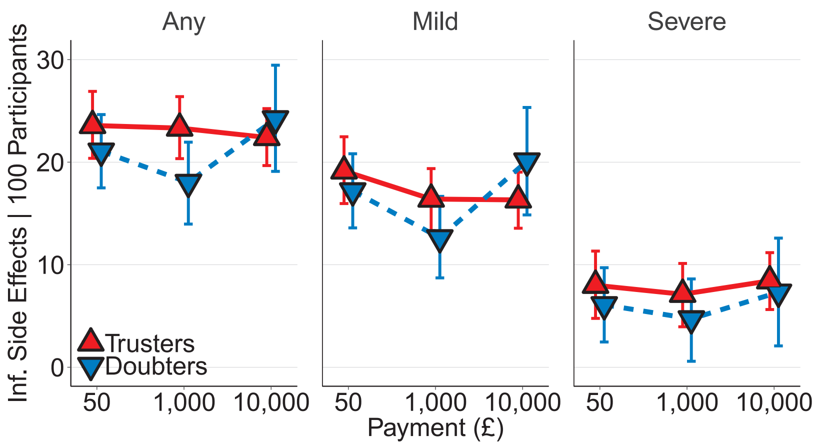


Figure 2. Estimated side effects for various payment amounts by Doubters and Trusters (between-respondents). Black triangles represent sample means. Colored triangles and error bars represent the means and the 95% highest density intervals of the posterior predictive distributions.

Specifically, as Figure 1 shows, Doubters considered the enrolment of a woman earning 1,500/month in the clinical trial as less voluntary when she was offered 10,000 ($b = -0.97$, $CI = [-1.49, -0.44]$). In addition, Doubters said a woman earning 1,500/month would be better off had she not even seen the high offer ($b = 1.26$, $CI = [0.26, 2.26]$). Consistent with this, Doubters also considered a woman earning 1,500/month to be more likely to regret accepting the offer ($b = 0.53$, $CI = [0.05, 1.02]$). Lastly, Doubters on an IRB panel approved the study offering 1,000 more than they approved the same study offering 10,000 ($b = -1.22$, $CI = [-1.69, -0.77]$).[4]

Inferred side effects

While Doubters found offering 10,000 to be less ethically appropriate than offering 1,000, Trusters found both amounts to be approximately equally appropriate. To investigate whether Doubters, unlike Trusters, employed high payment amounts as a cue to higher risks, we once again analyzed estimates for the first payment a respondent saw. We ran separate analyses for each type of side effect (any/mild/severe). There were no main effects of compensation amount on the estimated number of side effects (50 versus 1,000, as well as 1,000 versus 10,000; all CIs included 0). Moreover, as Figure 2 shows, increasing compensation from a very low amount (50) to a moderate amount (1,000) did not result in differential estimates of side effects for the two groups.

However, as in earlier analyses, individual differences emerged as a trial offered a very large compensation—10,000, compared to 1,000. Specifically, Doubters expected more side effects for trials that offered a very large compensation ($b_{any} = 6.19$, $CI = [1.15, 11.33]$, $b_{mild} = 7.47$, $CI = [3.30, 11.46]$, $b_{severe} = 3.00$, $CI = [-0.15, 6.17]$, note the CI includes 0 for severe side effects), whereas Trusters' inferred side effects

[4] All models tested the effect of offering 10,000 over 1,000, using "Truster" as a baseline. As in the original study (Ambuehl et al., 2015), the "IRB approval" variable is different from the variable used to define the types that appears later in the survey, in which all payment amounts were presented side by side, and participants were asked to judge the ethicality of each amount. However, the two variables are correlated: "IRB approval" (Figure 1) asked respondents to evaluate the trial *given* a particular payment. Later, respondents were asked to evaluate the trial and *determine* appropriate payment. As we show later, clinical trial evaluations can be explained using *inferred risks* and *voluntariness*—and without relying on the Truster/Doubter dichotomy.

were independent of payoff magnitudes (all CIs included 0).[5] When analyzing the estimated number of side effects within-respondents, Trusters also judged trials offering very high payments to be riskier, but they were less sensitive than Doubters to the payoff information in the vignette (Supplementary Material S3.2).

In sum, consistent with our hypothesis, we found a positive relationship between the amount a clinical trial offers and its inferred riskiness for Doubters but not Trusters. One qualification of this result is that extremely small payoffs (50) did not reduce inferred risks compared to moderate amounts (1,000)—only extremely large payoffs (10,000) increased inferred risks compared to moderate amounts (1,000).

Inferred side effects and clinical trial evaluations

Next, we examined whether respondents' ethical approval of the clinical trial was linked to their estimated number of side effects. Again, we focused on the first payment amount respondents saw. The judged number of participants to enrol in the study was unrelated to the estimated number of side effects (all CIs included 0): While the estimated number of side effects increased with increasing payment amounts from 1,000 to 10,000 (Figure 2), p *(enroll)* did not (Figure 1). It is plausible that inferred risk and high financial gain offset each other as respondents consider whether or not volunteers would enrol. However, a participant was considered to be better off without the offer if side effects were expected to be higher ($b_{any} = 0.034$, $CI = [0.025, 0.043]$). Consistent with this, respondents who estimated a higher number of side effects also thought that participants would be more likely to regret accepting an offer to take part in the study ($b_{any} = 0.022$, $CI = [0.017, 0.026]$). Lastly, higher risk assessments lowered respondents' IRB approval of the clinical trial ($b_{any} = -0.013$, $CI = [-0.017, -0.009]$, all modeled as main effects; coefficients are smaller for the side effects' regressor due to its scale $[0-100]$ as compared to the voluntariness scale $[0-7]$).

The role of estimated side effects in evaluating the trial did not depend on whether a respondent was a Doubter or a Truster (clinical trial evaluations modeled in three-way interaction using type [Doubter vs. Truster] × payment × estimated number of side effects as predictors; all CIs included 0). This is unsurprising since comparable interaction effects between these groups were present in clinical trial evaluations (Figure 1) and the estimated number of side effects (Figure 2).[6] These findings suggest that subjective risk estimates influence how ethical participants find a given trial.

Earlier research suggested that these evaluations were also related to how coercive respondents found the clinical trial (Ambuehl et al., 2015). To study whether both inferred risks and coercion (i.e., lower voluntariness) have separable influences on clinical trial evaluations, we simultaneously entered voluntariness and side effects as predictors in a linear regression; we found that both variables do explain unique variation in how respondents evaluated clinical trials. Participants were considered to be better off without seeing the offer when voluntariness was judged to be lower and inferred side effects were higher ($b_{volun.} = 0.3709$, $CI = [-0.4819, -0.2594]$, $b_{risk} = 0.0229$, $CI = [0.0209, 0.0390]$). Participants were

[5] All models tested the effect of offering 10,000 over 1,000 (or 1,000 over 50, using "Truster" as a baseline. Results of all analyses are listed in Supplementary Tables D3 and D4.

[6] The results also held in other specifications of the model, for instance after controlling for the [Doubter vs. Truster] × payment [50/1,000/10,000] interaction or the other side effect types as predictors [mild/severe]; see Supplementary Table D5.

thought to have less regret about accepting a trial if voluntariness was judged to be lower and inferred side effects were higher ($b_{volun.} = -0.2612$, $CI = [-0.3118, -0.2107]$, $b_{risk} = 0.0188$, $CI = [0.0146, 0.0229]$). Moreover, IRB approval increased when voluntariness was judged to be higher and inferred side effects were lower ($b_{volun.} = 0.2800$, $CI = [0.2291, 0.3311]$, $b_{risk} = -0.0098$, $CI = [-0.0139, -0.0057]$; regression tables reported in Supplementary Table D5).

We also explored the 'net effects' payoffs had on IRB approval for Doubters and Trusters using a mediation approach: As expected, there was a smaller net effect of payoff on IRB approval for Doubters and not Trusters after controlling for estimated side effects and coercion (Supplementary Material, Figure D1 and D2). A similar pattern emerged in within-respondent analyses: Doubters relied on payoff information when assessing risks and voluntariness, and ultimately when determining IRB approval. Trusters were more insensitive to changes in payoff magnitude.

Repugnance

In a final set of analyses, we also examined whether respondents who estimated a higher number of side effects also considered clinical trials as generally less ethically permissible. That is, can the perceived risk of a clinical trial at least partially explain whether or not individuals find clinical trials repugnant? We hypothesized that respondents who considered clinical trials to be riskier may also find them more repugnant, or less ethically permissible in general (independent of payoff magnitude). Indeed, higher repugnance was linked to a higher number of estimated side effects ($b_{10{,}000} = 0.0082$, $CI = [0.0058, 0.0106]$, estimates for 'any' side effects, given a payoff of 10,000). This link was stable across side effect estimates for different payment amounts and in other specifications of the model (Supplementary Table D6). The link between side effects and repugnance was present for Doubters ($b_{10{,}000} = 0.0082$, $CI = [0.0039, 0.0125)$ and Trusters ($b_{10{,}000} = 0.0048$, $CI = [0.0013, 0.0082]$; main effects; comparable results for 1,000 and 50)—again, the link was more pronounced for Doubters (comparison of β regressors).

Since coercion is another explanation for repugnance (Ambuehl et al., 2015), we also tested whether voluntariness affected repugnance ratings in a regression using both side effects estimates and voluntariness as predictors. Indeed, clinical trials were considered less repugnant when they were judged to be less coercive (or more voluntary). This held in addition to the variance explained by the estimated number of side effects ($b_{risk} = 0.0065$, $CI = [0.0042, 0.0088]$, $b_{volun.} = -0.1489$, $CI = [-0.1803, -0.1172]$, in a model relating repugnance to both side effects and voluntariness; see Supplementary Material D6) as well as to several demographic characteristics that can help explain who finds clinical trials repugnant, such as whether or not respondents had thought about participating in a clinical trial themselves (but not being a Doubter/Truster; see Supplementary Table D9 for full model).

General Discussion

More than 90% of healthy clinical trial volunteers say money is their main motivation for taking part (Bigorra and Baños, 1990; van Gelderen et al., 1993). However, the high pay—often significantly higher than other types of "unskilled" labor—these studies offer can make this transaction repugnant (Ambuehl et al., 2015; Roth, 2007; Wilkinson and Moore, 1997). There are at least two reasons that this is the case.

First, high payments can be seen as manipulative and even coercive. Second, as we hypothesized, high pay can be used as a cue to the risk involved. Both reasons can cause clinical trials offering high payments to be less ethically permissible.

Do people perceive higher payment to be associated with higher risk? We found this to be the case for one group of respondents. Doubters expected more people to experience side effects than people in the same trial with a smaller payoff. Trusters, in contrast, inferred similar risks across payoff magnitudes. Consistent with earlier results (Cryder et al., 2010), these findings do not support the conjecture that high pay *impairs* peoples' ability to think carefully about the risks and benefits of a clinical trial (Casarett et al., 2002); or, alternatively, that high pay conceals the risk involved in a clinical study (Bentley and Thacker, 2004; McNeill, 1997). To the extent that payments are actually correlated with risk (Grady, 2005), our results resonate with adaptive cognition theories (Anderson, 1991; Gibson, 1979; Gigerenzer et al., 1999; Marr, 1982; Simon, 1956; Perkovic and Orquin, 2017; Hertwig et al., 2013; Stewart, 2009). These theories state that the mind is often well-attuned to statistical structures in the environment: For instance, people are aware that the high rewards they desire are unlikely to occur due to fair-bet assumptions between seller and buyer (Pleskac and Hertwig, 2014). By the same token, people may also assume there are no "free lunches" in clinical trial markets and infer that high payments compensate for high risk.

Our analyses revealed individual differences in the extent to which inferred risks were payoff-dependent. There may be several reasons for this. First, some people may not consider payments and potential risks to be linked, either because the relationship is imperfect (in fact, only a third of surveyed research institutes claimed to factor in risk when determining payment; Grady, 2005), or because they have only been exposed to very few instances of the relationship (e.g., clinical trial advertisements). Second, since descriptions of clinical trials have to disclose known risks to potential participants, payoff information may not be the only cue people rely on to infer the riskiness of a clinical trial. For instance, in our vignette people were told that, *"Since no side effects occurred in animal studies, the institute's experts consider it unlikely that they will occur in humans."*, and *"If side effects occur, they may range from [...] nausea to [...] migraines".* The fewer additional cues they have, the more people may focus on payoff information (Leuker et al., 2018). More generally, people may have prior beliefs about the phase of the clinical trial and the medication in question. Thus, payment may be just one of several cues that could characterize a study's underlying risk. This may help explain variability in earlier results (Cryder et al., 2010; Mantzari et al., 2014).

From a purely economic perspective, high pay in clinical trials is not unethical. Why should someone be worse off if they are offered 10,000 instead of 1,000 for the same work? One concern—also voiced by our respondents—is that socioeconomically disadvantaged people in particular may feel coerced to take part (Ambuehl et al., 2015; Roth, 2007; Wilkinson and Moore, 1997). But another reason is that high payments are suspected to be associated with high risks. In contrast to a proposition by Ambuehl et al. (2015), Doubters made this inference for 1,000 versus 10,000 payment schemes despite the trials being otherwise identical. Ultimately such payoff-dependent expectations of more side effects from a very lucrative clinical trial can lead to lower IRB approvals, stemming from the expectation that more people would regret participating and the judgment that people would be better off without the offer. Higher pay does not monotonically increase IRB approval solely because it may be coercive, but also because it can change people's perception of risk.

In a final analysis, we showed that subjectively riskier trials were considered to be more repugnant. Objective risk has been established as playing a role in some repugnant markets (Table 1, p. 39, Roth, 2007), but to our knowledge the role of subjective, inferred risk has been overlooked when studying feelings of repugnance. This link between risk and repugnance may generalize to markets in which parties may be partially remunerated for the risk they take, such as surrogate motherhood.

Our findings may have practical implications for both research institutes and governments. Currently, extremely high payoffs are, by some, perceived to compensate for higher risks. However, severe incidents such as the 2015 French drug trial (AFP, 2016; Randerson, 2016) show that not even institutes themselves can accurately determine the risk a trial entails in advance. We suggest that research institutes should not compensate in advance for suspected higher risk. Instead, volunteers should be explicitly informed that they will be paid a fair wage for their time—and, as is already the case, all measures will be taken to minimize potential harm. If unanticipated side effects do occur, the research sponsors will compensate for research-related harm or injuries. This approach is consistent with Manning's (2017) statement that "the beneficiaries of research (researchers, sponsors, society) have a moral responsibility to compensate for research–related injury" (p. 426). This can be achieved, for instance, by implementing no-fault compensation (e.g., through a public fund). These policies may ultimately help reduce the repugnance of clinical trials.

References

AFP (2016). Man who died in French drug trial had 'unprecedented' reaction, say experts. *The Guardian*.

Ambuehl, S., Niederle, M., and Roth, A. E. (2015). More money, more problems? Can high pay be coercive and repugnant? *American Economic Review*, 105(5):357–360.

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, 98(3):409–429.

Bentley, J. P. and Thacker, P. G. (2004). The influence of risk and monetary payment on the research participation decision making process. *Journal of Medical Ethics*, 30(3):293–298.

Bigorra, J. and Baños, J. E. (1990). Weight of financial reward in the decision by medical students and experienced healthy volunteers to participate in clinical trials. *European Journal of Clinical Pharmacology*, 38(5):443–446.

Casarett, D., Karlawish, J., and Asch, D. A. (2002). Paying hypertension research subjects: Fair compensation or undue inducement? *Journal of General Internal Medicine*, 17(8):651–653.

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., and Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin Numeracy Test. *Judgment and Decision Making*, 7(1):25–47.

Cryder, C. E., John London, A., Volpp, K. G., and Loewenstein, G. (2010). Informative inducement: Study payment as a signal of risk. *Social Science and Medicine*, 70(3):455–464.

Dickert, N., Emanuel, E., and Grady, C. (2002). Paying research subjects: An analysis of current policies. *Annals of Internal Medicine*, 136(5):368–373.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.

Fric, C. (2016). Statutory minimum wages in the EU 2016.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472.

Gibson, J. (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston, MA.

Gigerenzer, G., Todd, P., and the ABC Research Group (1999). *Simple Heuristics That Make Us Smart*. Oxford University Press, New York, NY.

Grady, C. (2005). Payment of clinical research subjects. *The Journal of Clinical Investigation*, 115(7):1681–1687.

Hertwig, R., Hoffrage, U., and the ABC Research Group (2013). *Simple Heuristics in a Social World*. Oxford University Press.

Kruschke, J. K. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan, 2nd Ed.* Academic Press/Elsevier, New York, NY.

Leuker, C., Pachur, T., Hertwig, R., and Pleskac, T. J. (2018). Exploiting risk–reward structures in decision making under uncertainty. *Cognition*, 175:186–200.

Manning, J. M. (2017). Does the law on compensation for research-related injury in the UK, Australia, and New Zealand meet ethical requirements? *Medical Law Review*, 25(3):397–427.

Mantzari, E., Vogt, F., and Marteau, T. M. (2014). Does incentivising pill-taking 'crowd out' risk-information processing? Evidence from a web-based experiment. *Social Science and Medicine*, 106:75–82.

Marr, D. (1982). Vision: A computational approach.

McNeill, P. (1997). Paying people to participate in research: Why not? *Bioethics*, 11(5):390–396.

Nussbaum, M. C. (1998). "Whether from reason or prejudice": Taking money for bodily services. *The Journal of Legal Studies*, 27(S2):693–723.

Perkovic, S. and Orquin, J. L. (2017). Implicit statistical learning in real-world environments leads to ecologically rational decision making. *Psychological Science*, 29(1):34–44.

Pike, E. R. (2012). Recovering from research: A no-fault proposal to compensate injured research participants. *American Journal of Law & Medicine*, 38(1):7–62.

Pleskac, T. J. and Hertwig, R. (2014). Ecologically rational choice and the structure of the environment. *Journal of Experimental Psychology: General*, 143(5):2000–2019.

Randerson, J. (2016). Fatal French clinical trial failed to check data before raising drug dose. *Nature News*.

Resnik, D. B., Parasidis, E., Carroll, K., Evans, J. M., Pike, E. R., and Kissling, G. E. (2014). Related injury compensation policies of U.S. research institutions. *IRB: Ethics & Human Research*, 36(1):12–20.

Roth, A. E. (2007). Repugnance as a constraint on markets. *Journal of Economic Perspectives*, 21(3):37–58.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63(2):129–138.

Staff Writers (2018). National Minimum Wage. http://www.minimum-wage.co.uk/. Accessed: 2018-04-10.

Stan Development Team (2016). RStanArm: Bayesian applied regression modeling via Stan (version 2.9.0-4). R package version 2.13.1.

Stewart, N. (2009). Decision by sampling: The role of the decision environment in risky choice. *Quarterly Journal of Experimental Psychology*, 62(6):1041–1062.

van Gelderen, C., Savelkoul, T., van Dokkum, W., and Meulenbelt, J. (1993). Motives and perception of healthy volunteers who participate in experiments. *European Journal of Clinical Pharmacology*, 45(1):15–19.

Wilkinson, M. and Moore, A. (1997). Inducement in research. *Bioethics*, 11(5):373–389.

6 | Summary and Future Directions

"We demand rigidly defined areas of doubt and uncertainty!"

Douglas Adams, The Hitchhiker's Guide to the Galaxy

One of the big ideas in the study of decision making is that our preferences for any states of the world can be re–described in terms of (subjective) risks and rewards. Subjective risks and rewards can then be combined into an expected utility (Savage, 1954). Once these utilities are established, it is assumed that the decision maker can order all situations according to the strength of his or her preferences. Moreover, it is assumed that all choices are consistent with subjective utilities. Thus, at least in theory, how people should make decisions is fairly straightforward.

Unfortunately, many real–world decisions involve grappling with uncertainty, where probabilities are not given or are difficult to ascertain (epistemic uncertainty). In other decisions, probabilities are given but randomness is not fully removed from the world, i.e. the probabilities are not 0 or 1 (aleatory uncertainty). Another complication is that environments often change (dynamic uncertainty), which makes it hard to predict the future from the present and hard to know which decision strategy to use in which situation (systemic uncertainty). Thus, while in theory decision making is easy, different types of uncertainty in the world create "stumbling blocks". A big stumbling block is the question of where probabilities come from in decisions under uncertainty (Meder et al., 2013). Another stumbling block is how much decision makers should sample in a given choice environment (even if probabilities are, in principle, available), or how they can anticipate rare events.

It seems that if a decision maker wants to choose *as if* he or she maximized subjective utilities, it is up to him or her to rigidly define areas of doubt and uncertainty. One way to achieve this may be to exploit the relationship between risks and rewards that exists in many monetary and non-monetary domains in the environment. In this dissertation, I theoretically and empirically examined how the mind adapts to risk–reward structures and how the link between risks and rewards impacts judgments and decisions. The basis of these investigations were theories of adaptive cognition, according to which the environment is represented or reflected in the mind and these representations systematically affect how it operates (e.g., Anderson and Schooler, 1991; Brunswik, 1944; Gibson, 1979).

In chapters 2, 3 and 4, these questions were investigated experimentally by—as many researchers have done before—turning to the world of monetary gambles. In these experiments, I attempted to extend the scope of "small worlds" (Savage, 1954) to slightly "larger worlds", by characterizing choices

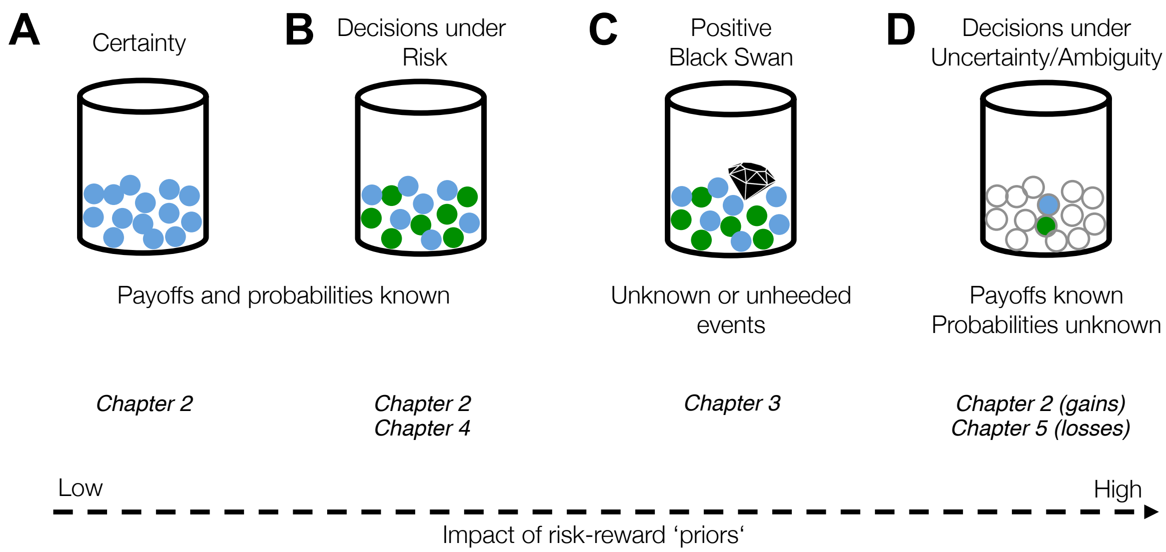


Figure 1. Decision making scenarios × risk–reward structures investigated in this dissertation. **(A)** Decisions under certainty were unaffected by risk-reward structure. When we interspersed gambles that are known to produce the 'certainty effect' in environments with different risk–reward environments, we found a standard certainty effect in all of these environments. **(B)** Chapter 4 showed that prior exposure to a negative risk–reward structure elicits more frugal processing strategies than prior exposure to uncorrelated risk-reward relationships in decisions under risk. Moreover, probabilities in decisions under risk were never completely ignored (i.e. no purely noncompensatory processing). This is consistent across all experiments in which probability information was presented. Despite changes in processing, there were no changes in risk preferences when modeling choices in a Prospect Theory framework. These modeling results are consistent across experiments (Experiment 1 & 2 in Chapter 2, Experiment in Chapter 4). **(C)** Risk-reward structures are a context variable that determines which values people can expect in a given environment. Surprising options (which contradicted the risk–reward context) were evaluated with more scrutiny, especially high payoff/high probability options, but there was no reliable evidence for a change in preferences depending on whether an option was surprising or expected. In the uncertainty framework, surprising options can be best described as positive black swans. **(D)** Chapter 2 shows that learned risk-reward structures are exploited in decisions under uncertainty via use of the risk-reward heuristic, according to which probabilities can simply be inferred from payoffs, ultimately resulting in environment-contingent preferences. In this framework, clinical trials can be conceptualized as uncertain environments in which the uncertainty pertains to the (often unknown) probability of side effects occurring. Assumed risk-reward structures can affect judgments in this context such that high payoffs make clinical trials appear riskier (Chapter 5).

across environments in which I systematically varied the overall risk–reward structures that generated the gambles. Across all experiments, participants were sensitive to the environments used to generate gambles. The extent to which risk–reward structures impacted judgments and decisions varied, among other variables, depending on the level of uncertainty involved in the choice task (for an overview see Figure 1).

In Chapter 2, I showed how people can exploit risk–reward structures in decisions under uncertainty by using a risk–reward heuristic. Consistent with the principle of *vicarious functioning* of different cues (Brunswik, 1943), people can (and did) infer the probabilities directly from the payoffs. The experimental data was consistent with two key requirements of a risk–reward heuristic, (1) people are able to extract the environmental structure and (2) people seem to adaptively use this structure as the ecological regularities vary across environments. Across three experiments, I demonstrated that people can learn risk–reward relationships from the options they experience during preferential choice. Participants learned the structures in an unsupervised, incidental manner—without being asked to explicitly attend to these structures. The learned risk–reward relationships guided inferred probabilities and ultimately impacted peoples' preferences in decisions under uncertainty.

Chapter 3 showed that being exposed to a particular risk–reward structure leads people to build expectations about options in subsequent trials. People's evaluations of risky options are based not only on the options' payoffs and probabilities. Part of them are also based on the extent to which they fit the risk–reward structure of the environment. These effects of "surprise" were traced in longer response times and an increase in pupil size, in particular when options were "surprisingly good"—i.e., they offered a high payoff and a high probability in an environment in which risks and rewards were otherwise inversely related. This may reflect how people process options in non-laboratory environments, in which high payoff/high probability options are often "too good to be true" or even positive black swans (Figure 1C). If you ever got an email from a Nigerian prince offering you a generous 20% share of $1,000,000,000 to help transfer the amount out of Nigeria to your account, it is probably a scam. Similarly, if an editor of a journal approaches you to publish your research immediately and without peer review, it too is probably a scam. The way we realize these are scams is because they violate the risk–reward structure we have learned.

Chapter 4 is also concerned with decisions under risk, but characterizes how learning about one of three risk–reward environments impacts preferences and evidence accumulation without any surprising options appearing (Figure 1B). As risks and rewards have by and large been treated as uncorrelated in empirical studies of risky choice, theories derived from them may only captured choice behavior in a fraction of the environments people typically experience outside the lab. Indeed, a computational model of eye tracking and behavioral data revealed that risk–reward structures can affect how they process options in decisions under risk. Specifically, inversely related risks and rewards can result in lower response thresholds and a higher gaze bias in the evidence accumulation process. Conversely, uncorrelated risks and rewards lead to more rigorous processing of options in risky choice and a relatively equal distributions of gaze across attributes and alternatives. In sum, uncorrelated risk–reward environments can promote maximizing—that is, more rigorous processing and closer-to-normative choices; while environments in which risks and rewards are inversely related allow, and tend to promote, satisficing. These results did not generalize to (1) conditions where participants learned a risk–reward structure explicitly from feedback, (2) conditions with very strong time pressure (1.5s response deadline), (3) learning about a counter-intuitive, positive risk–reward relationship. Moreover, this Chapter—together with evidence from Chapter 2—revealed that risk–reward structures do not systematically affect participants' risk preferences. For instance, we found evidence for the certainty effect in both negative and uncorrelated risk–reward environments in Chapter 2. Moreover, subjective evaluations of risks and rewards as estimated via a Prospect Theory framework were comparable across risk–reward environments.

In chapter 5, I showed people's perceptions of market mechanisms—specifically, their intuition that there usually is "no free lunch"—affects inferences regarding some of the options people may be offered outside the lab. Specifically, survey respondents evaluated a medical trial for a new Ebola vaccine offering high or low payment amounts. The survey revealed that some individuals used very high pay as a cue to infer that the trial must pose a high risk. Higher risk assessments were were linked to lower ethicality assessments of the clinical trial. Clinical trials can be best conceptualized as decisions under uncertainty with unknown probabilities (Figure 1B), especially if they are phase I clinical trials—that is, trials that move from animal testing to humans. Although as much uncertainty as possible has been reduced in pretests with animals, researchers can never know for sure how humans will react to certain substances. This level of uncertainty also exists for potential volunteers who review clinical trial advertisements.

Synthesis

This dissertation identified some general factors that determine the extent to which people rely (purely) on a risk–reward heuristic. It seems that the more the environment allows people to reduce uncertainty, for example by gathering or using other information, the less people rely on the risk–reward heuristic. Chapter 2 (experiments 1 & 2) showed that people strongly rely on it when no other information but the payoff is given to them—i.e., in decisions under (complete) uncertainty. When *some* context is available and people have to gauge their chances of winning based on epistemic events (such as in soccer bets or as in Chapter 2, previous expectations about the weather), both the epistemic information and the inferences from the risk–reward heuristic are integrated into a common probability estimate. Interestingly, even within this experiment, the risk–reward relationship affected probability estimates most when the probabilities associated with the epistemic event hinged at 50%—that is, when uncertainty (or entropy) was highest (this result is shown in the appendix to Chapter 2). Eye tracking data in Chapter 3 and Chapter 4 showed that when the probability information is available in descriptive form, as is the case in decisions under risk, people are likely to use it (instead of heuristically inferring it). It can make good ecological sense to have keep track of explicit probability information if it is available—as it would help to detect above–average options that decision makers would like to exploit (e.g. when "beating the bookies", see Kaunitz et al., 2017), and discard below–average, dominated options they would like to avoid.

Across chapters, some general factors that determine how (well) people learn risk–reward structures emerged. Most of the presented studies focused on incidental learning, which reflects how people often learn in non-laboratory environments, where people often learn without the luxury of explicit feedback. In the case of decision making, learning is also not the primary goal; typically, the aim is to find the most promising alternative in a set of options. Chapter 4 also showed that incidental learning through evaluating the options is somewhat special: it allows people to learn the risk–reward relationships but also the range of possible values (EVs) to be expected in a given environment. Chapters 2–4 did not include a (control) condition in which participants made decisions under risk or uncertainty without learning. Do risk–reward structures impact decisions without any experimentally induced learning? Two sets of results speak for this conjecture. First, in earlier studies, participants were asked for their probability estimates given different payoff magnitudes (Pleskac and Hertwig, 2014) or for their payoff estimates given different probability magnitudes (Skylark and Prabhu-Naik, 2018). In these cases, participants assumed risks and rewards to be inversely related, suggesting that the inverse relationship between risks and rewards is frequent and recurrent. Second, Chapter 5 showed that people assume that "there is no free lunch" in clinical trials—and inferred that high pay insinuated high risk. Notably, this was the case without prior learning and perhaps even without a link between high pay and high risk consistently present in the domain of clinical trials. Lastly, across all Chapters, when learning incidentally, the less natural positive risk–reward relationship was harder to learn. This again speaks for the pervasiveness of the inverse risk–reward relationship.

The findings presented in this dissertation have implications for theories of decision making. I showed that the mind is prepared to learn about risk–reward structures from the world, and also flexibly adapts to them as they change. By providing a theory about how probability estimates are generated, it helps to overcome one of the big "stumbling blocks" of making decisions—namely that it is often unclear where probabilities come from. People are commonly thought to intuit subjective probabilities by estimating statistical probabilities from samples of information (Hertwig and Erev, 2009), or from their knowledge and memory (Fox and Tversky, 1998). The results presented in this dissertation suggest another ecologically-grounded solution: namely, that people estimate the missing probabilities from their immediate choice environments via their learned risk–reward relationships. Choices are consistent with these estimates, meaning that preferences also change based on recently experienced environmental structures. These results support the notion that preferences are often constructed from local choice contexts rather than revealed (Lichtenstein and Slovic, 2006). It is fair to say that forming preferences within a local choice environment this way is ecologically rational. More broadly, the notion that preferences are at least partially constructed given local choice contexts may help explain limited test–retest reliability in behavioral measures of risk preference (Frey et al., 2017).

The idea that payoffs impact risk estimates in decisions under uncertainty also has practical implications for research institutes who wish to recruit volunteers. First, higher pay does not mean higher participation rates—as long as (some) people equate higher pay with higher risk. In addition, true underlying risks hard to be determined in the first place which would render the link (if it were desired) unreliable. Consequently, as proposed in chapter 5, clinical trial researchers should communicate that payment is not a signal for the underlying risk, but reflects a compensation for time and inconvenience. If side effects are experienced, states should offer a no–fault compensation to any affected volunteer. Assumed links between payment and risk may also partially explain "motivational crowding out" effects, in which price incentives sometimes undermine an individual's moral obligation—e.g. to donate blood, or on a societal level to store nuclear waste. Monetary incentives have been suggested to reduce intrinsic motivation, and to some, compensation amounts may also signal underlying risks (Frey and Oberholzer-Gee, 1997).

Outlook

Several questions newly emerged or were left unanswered in this dissertation. A first question is how the results presented here generalize across different domains of losses. Chapter 5 shows how (assumed) risk–reward structures impact a particular loss domain, namely side effects experienced in a clinical trial. On the one hand, these results suggest that people may apply a similar heuristic—one that weighs expected gains and expected losses—across gains and losses and different magnitudes thereof: In previously studied domains such as monetary lotteries, the loss is usually small and inevitable (for instance, a pay–to–play fee of 2$). In clinical trials, the loss is larger but rarer. Yet the mechanisms that relates gains, probabilities, and losses may be similar, at least in the context of markets where prices are pushed towards "fair" bets. Here, a transaction or bet is fair when the gain g offsets the expected loss c:

$$p \times g = (1 - p) \times c.$$

In a fair bet, the expected net gain should be zero. Thus, the probability of a gain can be inferred by solving for p:

$$p = c/(c + g)$$

At the same time, choices in the loss domain may be influenced by many other aspects such as affective reactions (Finucane et al., 2000), loss aversion (Tversky and Kahneman, 1992), (in-)experience with situations in which losses are incurred, or beliefs about the event occurring (especially if they are rare events), which often influences insurance choices (Slovic et al., 1977). Therefore, the extent to which probabilities can be inferred from the magnitude of the loss may be domain–specific (Pleskac and Hertwig, 2014). An ecological analysis of risk–loss relationships in domains of interest could shed light on this question.

Second, individuals vary in the extent to which they apply the risk–reward heuristic in decisions under uncertainty (Chapter 1) or in the domain of losses (Chapter 5). It is unknown what discerns between people who rely more on a risk–reward heuristic or less on it. It is plausible that this related back to how quickly individuals *can* pick up statistical regularities from their environment. In this case, older adults may be less flexible in adapting to new environments due to cognitive decline—in particular due to decreases in fluid intelligence and a potentially stronger reliance on pragmatics, or crystallized intelligence (Mata et al., 2015; Baltes et al., 1999). It could also be that some people have stronger priors, and thus update their beliefs based on environments they encounter less flexibly (e.g., Chater et al., 2010). It may be an adaptive strategy to not immediately update one's beliefs after seeing options that are too good to be true; instead, it may be an adaptive strategy to follow a "rule-plus-exception model," according to which people may learn exceptions to a rule instead of updating that rule if they are unable to identify the alternative rule that would account for the exceptions (Nosofsky et al., 1994). Lastly, people may not be equally exposed to different environments. For instance, people may not have come across clinical trial advertisements very frequently, or the relationship between payoffs and risks may be an unreliable one.

A third question pertains to how risk–reward structures affect decisions from experience (Hertwig and Erev, 2009). Results from Chapter 4 suggest that people may seek for more evidence in uncorrelated environments than in environments in which risk–reward structures are consistent with their priors from non-laboratory environments. Some results consistent with this conjecture have been obtained by other authors who studied participants' sampling behavior in an experience–based probability judgment task (Rieskamp et al., 2017). Sampling paradigms can also be used to address how commonly different payoff magnitudes are experienced. There is an interesting discrepancy between the payoff magnitudes that people hear about in the media—which may be given greater weight by the mere fact that they are widely discussed (for instance as people winning lottery jackpots are reported, or the salaries of famous soccer players)—, and the typically smaller payoffs people incidentally experience (Ungemach et al., 2011).

Can the mind ever predict what risk–reward relationship will be encountered in a particular environment, without substantial experience in the environment? In recent work, Pleskac et al. (prep) have looked into a deductive approach to predicting risk–reward environments that is based on the theory of ideal free distribution. Based on this work, a negative risk–reward structure is to be expected, as the number of competitors will typically be high for large resource items. However, a degradation of the risk–reward relationship is predicted when the system falls out of equilibrium (e.g. when the number of competitors is low or when they are not "ideal" agents). I think this work is exciting: If the mind has a good (deductive) theory of the environmental structure, there is hope that decision-makers can minimize systemic uncertainty (arising from the mind) across many domains.

References

Anderson, J. R. and Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 2(6):396–408.

Baltes, P. B., Staudinger, U. M., and Lindenberger, U. (1999). Lifespan Psychology: Theory and Application to Intellectual Functioning. *Annual Review of Psychology*, 50:471–507.

Brunswik, E. (1943). Organismic achievement and environmental probability. *Psychological Review*, 50(3):255–272.

Brunswik, E. (1944). Distal focussing of perception: Size constancy in a representative sample of situations. *Psychological Monographs*, 56(254):1–49.

Chater, N., Oaksford, M., Hahn, U., and Heit, E. (2010). Bayesian models of cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(6):811–823.

Finucane, M. L., Alhakami, A., Slovic, P., and Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, 13(1):1–17.

Fox, C. R. and Tversky, A. (1998). A belief-based account of decision under uncertainty. *Management Science*, 44(7):879–895.

Frey, B. S. and Oberholzer-Gee, F. (1997). The cost of price incentives: An empirical analysis of motivation crowding–out. *The American Economic Review*, 87(4):746–755.

Frey, R., Pedroni, A., Mata, R., Rieskamp, J., and Hertwig, R. (2017). Risk preference shares the psychometric structure of major psychological traits. *Science Advances*, 3(10).

Gibson, J. (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston, MA.

Hertwig, R. and Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences*, 13(12):517–523.

Kaunitz, L., Zhong, S., and Kreiner, J. (2017). Beating the bookies with their own numbers-and how the online sports betting market is rigged. *arXiv preprint arXiv:1710.02824*.

Lichtenstein, S. and Slovic, P. (2006). *The Construction of Preferences*. Cambridge University Press, New York, NY.

Mata, R., Josef, A. K., and Lemaire, P. (2015). Adaptive Decision Making and Aging. In *Aging and Decision Making: Empirical and Applied Perspectives*, pages 105–126.

Meder, B., Lec, F. L., and Osman, M. (2013). Decision making in uncertain times: What can cognitive and decision sciences say about or learn from economic crises? *Trends in Cognitive Sciences*, 17(6):257–260.

Nosofsky, R. M., Palmeri, T. J., and McKinley, S. C. (1994). Rule-plus-exception model of classification learning. 101(1):53–79.

Pleskac, T. J., Conradt, L., Leuker, C., and Hertwig, R. (in prep). The ecology of competition: A linchpin for risk–reward structures. *Manuscript in prepartion*.

Pleskac, T. J. and Hertwig, R. (2014). Ecologically rational choice and the structure of the environment. *Journal of Experimental Psychology: General*, 143(5):2000–2019.

Rieskamp, J., Hoffart, J., and Dutilh, G. (2017). How the environment influences information search in decision making. Talk at the 26th Subjective Probability, Utility and Decision Making Conference (SPUDM).

Savage, L. J. (1954). *The Foundations of Statistics*. John Wiley & Sons, New York, NY.

Skylark, W. J. and Prabhu-Naik, S. (2018). A new test of the risk-reward heuristic. *Judgment and Decision Making*, 13(1):73.

Slovic, P., Fischhoff, B., Lichtenstein, S., Corrigan, B., and Combs, B. (1977). Preference for Insuring against Probable Small Losses: Insurance Implications. *The Journal of Risk and Insurance*, 44(2):237–258.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323.

Ungemach, C., Stewart, N., and Reimers, S. (2011). How incidental values from the environment affect decisions about money, risk, and delay. *Psychological Science*, 22(2):253–260.

Appendices

A | Supplementary Material to Chapter 2: "Exploiting risk–reward structures in decision making under uncertainty"

1. Experiment 1

Learning phase (detailed methods)

For the negative condition, the gambles were constructed as follows. Payoffs were 100 random draws from a uniform distribution with a range 1.01–1000. The probabilities for each payoff were set so that they were inversely related to the payoff x: $p = 1 - x/1000$. We then jittered payoffs and probabilities by adding normally distributed noise with a standard deviation of 0.2 to the logit transformation of the probabilities and to the logit transformation of normalized payoffs, and we transformed those perturbed values back to the scales used in the experiment. We paired the 100 initially created gambles to all possible unique combinations (4,950), and from all possible combinations randomly sampled 100 nondominated pairs for use in the experiment. Differences in expected value between gambles in this condition were relatively small (Mdn_{abs} = E\$67.7, $Mdn_{\%}$ = 6.8%, range E\$0.17–E\$281.87).

For the uncorrelated condition, we took the 100 gamble pairs used in the negative condition, but now randomly linked probabilities and payoffs. We did this to maintain the marginal distributions of payoffs and probabilities across both the conditions and the total number of gambles in the set (see Stewart et al., 2006, 2015). We again drew 100 random gamble pairs from all possible combinations. If any of the gamble pairs had stochastically dominated options (i.e., $p_A > p_B$ and $x_A > x_B$), we switched the probabilities of gambles A and B. The expected value differences between gambles were larger in the uncorrelated condition than in the negative condition (Mdn_{abs} = E\$133, $Mdn_{\%}$ = 13.3%, range E\$0.52–E\$844.6).

Experimental setup

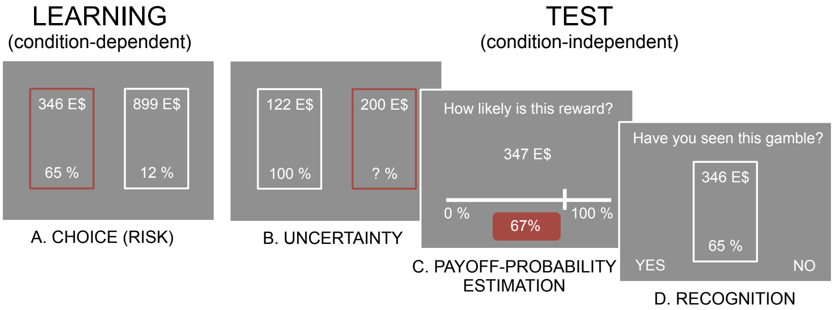


Figure A1. Tasks in Experiment 1. (A) Participants completed a learning phase in which they chose between nondominated gamble pairs drawn from one of two risk–reward conditions (between–participants, 119 trials). (B) Then, all participants chose between an uncertain option and an approximately half-as-large certain option (20 trials). The task also included 20 filler trials, in which the certain options were created by using different fractions of the uncertain option. (C) Participants estimated the probabilities associated with 10 different payoffs. (D) Participants completed a recognition task pertaining to gambles in the main experiment (16 targets, 16 foils). The location of payoffs and probabilities (top/bottom) was randomized between participants (but same location across all tasks). The order of (C) and (D) was randomized between participants.

Processing in Decisions under Risk: Certainty effect

Hypotheses

In correlated risk–reward environments such as the negative condition in Experiment 1, one attribute (the payoff, or the probability) is, in principle, redundant as it can be inferred from the other. Thus, one could in principle employ a noncompensatory processing strategy. We used gambles that are known to produce the certainty effect to test for such noncompensatory processing strategies in the negative vs. uncorrelated risk–reward environment. We hypothesized that the certainty effect would be reduced in the negative condition, as people might attend less to the probability information (i.e. "100%").

			Proportion choosing A	
	Choice	Gamble Pair	Negative	Uncorrelated
Low Payoff	1	**A** $(96, 1.0; 0)$ or **B** $(128, .80; 0)$	**.75**	**.83**
	2	**A** $(96, .25; 0)$ or **B** $(128, .20; 0)$	.38	.33
High Payoff	3	**A** $(746, 1.0; 0)$ or **B** $(995, .80; 0)$	**.84**	**.80**
	4	**A** $(746, .25; 0)$ or **B** $(995, .20; 0)$	.30	.28

Table A1. Certainty effect gambles used in learning phase of experiment 1 and proportion of participants choosing gamble A in the negative and uncorrelated condition. The gambles 'x with a probability p otherwise 0' are represented in the table as $(x, p; 0)$. Choice proportions are aggregated across participants within each condition.

Results

Did participants in the negative condition change to using noncompensatory processing strategies, ultimately leading to a smaller certainty effect? We did not find evidence for this. As in the standard certainty effect, most participants preferred the certain but lower payoff option to a risky but higher payoff (Table A1, $M_{\text{choose(A)}} = .81$, $b = 2.05$, $CI = [1.13, 3.10]$). These preferences switched when payoffs remain the same but probabilities were scaled down ($M_{\text{choose(A)}} = .33$, $b = -3.03$, $CI = [-3.97, -2.20]$). There were no credible effects of condition ($b_{\text{negative}} = -0.02$, $CI = [-1.06, 0.99]$), or payoff magnitude ($b_{\text{high}} = 0.11$, $CI = [-0.76, 0.55]$) on the size of the certainty effect (logistic regression using probability level, payoff level and condition as predictors; and participant as a grouping factor).

As the certainty effect is explained by the inverse s-shaped probability weighting function in prospect theory, typically understood as overweighting of rare events, this suggests little difference in how participants subjectively valued payoffs and probabilities between the two conditions. Indeed, when we fit prospect theory to the choice data, we found very little difference between how participants in the two conditions subjectively weighed payoffs and probabilities (see below).

Prospect Theory Model

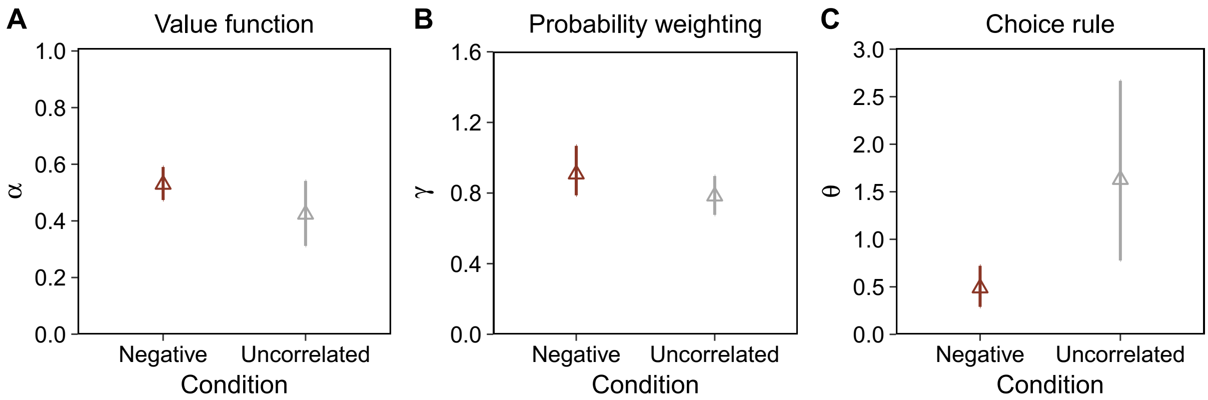


Figure A2. Parameters from Prospect Theory Model for Study 1. We implemented a Bayesian Hierarchical Model. Plots show the group hypermeans and 95% Highest Density Intervals (HDI) per condition, for the value function parameter (A), the probability weighting function (B) and the error parameter (C).

Model Specification

We used all 119 gambles per condition of which 100 gambles were condition-dependent learning trials and 19 gambles were common between conditions. We implemented a Bayesian Hierarchical Version of Prospect Theory using JAGS to sample from posterior distributions. Our priors were comparable to those implemented in the Cumulative Prospect Theory Model in Scheibehenne and Pachur (2015), except that priors for alpha and gamma ranged from 0-2 (and were not restricted between 0 and 1).

The Bayesian Hierarchical approach allowed us to fit the model at the participant level and at the group level simultaneously. We used the a standard power function for the value function, a one-parameter probability weighting function (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), and a Luce choice rule as an error theory (Luce, 1959; McFadden, 1976).

Model Interpretation

Values closer to 1 in the value and the probability weighting function indicate less subjective distortions in values or probabilities. We do not find credible differences in these parameters across conditions, suggesting that participants do not treat payoff or probability information very differently across conditions.

When the sensitivity parameter θ is 0 this is equal to random choice (choosing either gamble with $p = .5$). When it increases choice becomes more deterministic and is guided increasingly by the expected utility differences between the two gambles. This is what we observe for the uncorrelated condition (see Figure A2). By design of the experiment we did not control for expected value differences in the learning trials. The uncorrelated condition is characterized by larger expected value differences whereas the negative condition is characterized by smaller expected value differences (payoffs and probabilities always trade off against one another). These differences in gamble pairs likely produced the differences in the estimated choice rule parameter, θ.

2. Experiment 2

Learning phase (detailed methods)

The methods were largely the same as in Experiment 1: Payoffs were 150 random draws from a uniform distribution with a range 1.01–2500. The probabilities for each payoff were set so that they were inversely related to the payoff x: $p = 1 - x/2500$. As can be seen from the formula, we used a larger payoff range ($E\$1.01 - 2500$, disclosed conversion rate $E\$2500 =$ EUR1). To create the positive risk–reward condition, we took the gambles in the negative condition and reversed the order of probabilities such that the highest probabilities were now associated with the highest payoffs (and vice versa). To create the uncorrelated condition, we randomly paired payoffs and probabilities. As before, this procedure was used to maintain the same marginal distributions of payoffs and probabilities across conditions.

As a common set of gambles, we included 10 gambles in the center of the payoff–probability distribution space (intermediate payoffs and probabilities) and 12 gambles at the margins of that space (3 high payoff/high probability, 3 high payoff/low probability, 3 low payoff/high probability, 3 low payoff/low probability; see triangles in Figure 4). Payoffs were random draws between $E\$0 - 500$ (low) and $E\$2000 - 2500$ (high). Probabilities were random draws between $0 - .2$ (low) and $.8 - 1.0$ (high). As in Experiment 1, these gambles were used to study whether there were any condition-dependent differences in how participants priced the gambles, while controlling for crucial factors such as EV differences between gambles. In total, these procedures resulted in 172 risky gambles per risk–reward condition while controlling for the marginal distribution of payoffs and probabilities across all three conditions.

The way the environment gambles were generated resulted in different distributions of expected values in the learning phase, although the marginal distributions were maintained ($Md_{\text{negative}} = 501.9$E\$, range $= 3.0 - 702.1$E\$, $Md_{\text{positive}} = 598.2$E\$, range $= 0.03 - 2485.0E\$$, $Md_{\text{uncorrelated}} = 507.1E\$$, range $= 0.24 - 2250.4E\$$).

Experimental setup

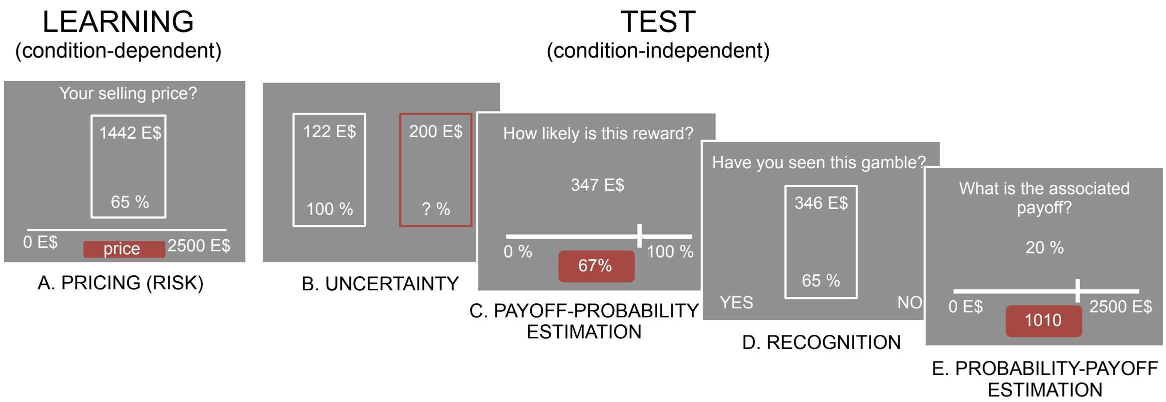


Figure A3. Tasks in Experiment 2. (A) Participants completed a learning phase in which they priced one gamble at a time drawn from one of three risk–reward conditions. (B) Then, all participants chose between an uncertain option and an approximately half-as-large certain option (20 trials). The task also included 20 filler trials, in which the certain options were created by using different fractions of the uncertain option. (C) Participants estimated the probabilities associated with 20 different payoffs. (D) Participants completed a recognition task pertaining to gambles in the main experiment. (E) Participants estimated the payoffs associated with 20 different probabilities. The location of payoffs and probabilities was randomized between participants.

Prospect Theory Model

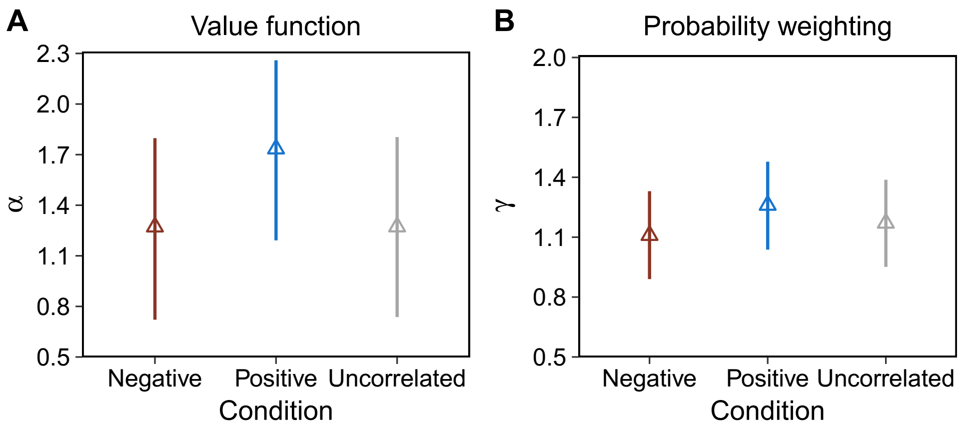


Figure A4. Parameter estimates from Prospect Theory Model for Study 2. Plots show the means and 95% Highest Density Intervals (HDI) per condition, for the value function parameter (A) and the probability weighting function (B).

Model Specification

We used all 172 learning phase trials to fit a Prospect Theory model from the certainty equivalents participants gave. We implemented a regression model that allowed us to estimate value function and probability weighting function parameters simultaneously for all participants in the experiment (Prelec, 1998). Specifically, α and γ can be estimated by

$$-ln(-ln(\frac{certainty\ equivalent}{payoff})) = ln\ \gamma + \ \alpha(-ln(-ln\ probability))$$

We used this model to run a regression that allowed for individual variation in both the slope and the intercept. The slope term can be interpreted as an estimate of γ, the back–transformed intercept term can be interpreted as an estimate of α. This model does not have an explicit error theory (errors can be conceptualized as residuals in the regression).

Model Interpretation

Similar to Experiment 1, there were no meaningful differences in average estimated value function (α) and probability weighting function parameters (γ) across different risk–reward conditions. However, both parameters exceed 1. How should this be interpreted? Value function parameters above 1 are typically interpreted as "risk–seeking". The task in Experiment 2 was to indicate a willingness to sell for gambles (rather than a willingness to accept), which means that participants were endowed with a gamble and had to indicate a price they would be willing to sell a gamble for.

Thus, it is plausible that high prices and consequently high α values were driven by an endowment effect. This is consistent with a recent meta–analysis that finds alpha values exceeding 1 in a majority of studies asking participants to indicate a willingness to sell (Yechiam et al., 2017), but not a willingness to buy.

The probability weighting function parameters, γ, slightly exceeding 1 suggest an s-shaped probability weighting function. The fact that values were not very far off 1 is partially consistent with Yechiam and colleagues' findings in that willingness to sell paradigms often produced less "biased" probability weighting functions (i.e. close to 1) than willingness to pay paradigms. Here we did not directly compare the two paradigms, so we can only speculate whether the results would have been different if we had tasked people to indicate a willingness to pay.

Regarding absolute γ values (exceeding 1 across conditions), our results deviate from Yechiam et al.'s in that mean parameters were often smaller than 1. There are two points to be made here. First, the credible intervals of estimated probability weighting parameters also contained values above 1 in some of the experiments, effectively rendering such values plausible estimates. Second, Yechiam et al. (2017) implemented a response bias parameter β in their model, and found a relatively larger response bias for willingness to sell versus willingness to buy decisions. Such a response bias may have been captured by γ in our model.

Estimating payoffs from probabilities (payoff–probability estimation task)

At the end of Experiment 2, we asked participants to estimate payoffs from probabilities. As Figure A5 shows, the estimates again reflected the conditions in the learning phase. Thus, both payoffs but also probabilities can serve as cues in risk–reward tasks, when either information is missing.

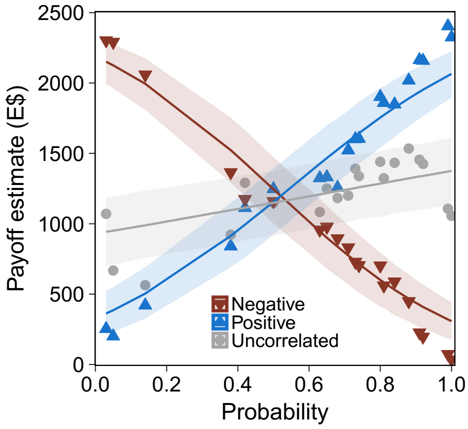


Figure A5. Payoff estimation task results by condition. Triangles show mean estimates by condition and probability magnitude. Lines and ribbons indicate the mean [95% HDI] of the posterior predictive intervals. Results were modeled with normalized reward estimates [0,1] on the logit scale such that (backtransformed) posterior predictions were confined between [0,2500]. The results reflect the risk-reward relationship each condition's participants had been exposed to, and are consistent with data from the probability estimation task.

3. Experiment 3

Learning phase (detailed methods)

During the learning phase, participants priced gambles based on the Berlin weather. To construct these gambles, we retrieved past weather data on the mean ($M = 22.7°C$) and standard deviation ($SD = 3.2°C$) of the maximum daily temperature in August in Berlin in 2011 from *accuweather.com*. We created 155 temperature ranges of varying width and location on the temperature scale (each August date from $1st$ – $31st$ was used 5 times) (see Figure A6). Because the maximum temperatures were approximately normally distributed, we calculated the historical frequency to approximate the probability that the maximum temperature on a given date would fall within the specified interval. As before, we then constructed gambles such that there was either a positive or a negative risk–reward relationship holding the marginal distributions of payoffs and historical frequencies constant across conditions (see Figure A7).

In the learning under uncertainty condition, only the temperture range and the corresponding payoff was shown for each gamble (e.g., "E\$2300 if the maximum temperature was between 13 and 15$°C$ on Aug 29th".) In the learning under risk condition, the historical (relative) frequency was added to the gamble (e.g., "E\$2300 if the maximum temperature was between 13 and 15$°C$ on Aug 29th (p = 3%)."). For Screenshots see Figure A8. As in the other experiments, the resulting "EV" distributions differed across the two conditions (Md_{negative} = 410.45E\$, range = 21.12 – 660.9E\$, Md_{positive} = 540.33E\$, range = 0.88 – 2355.8E\$).

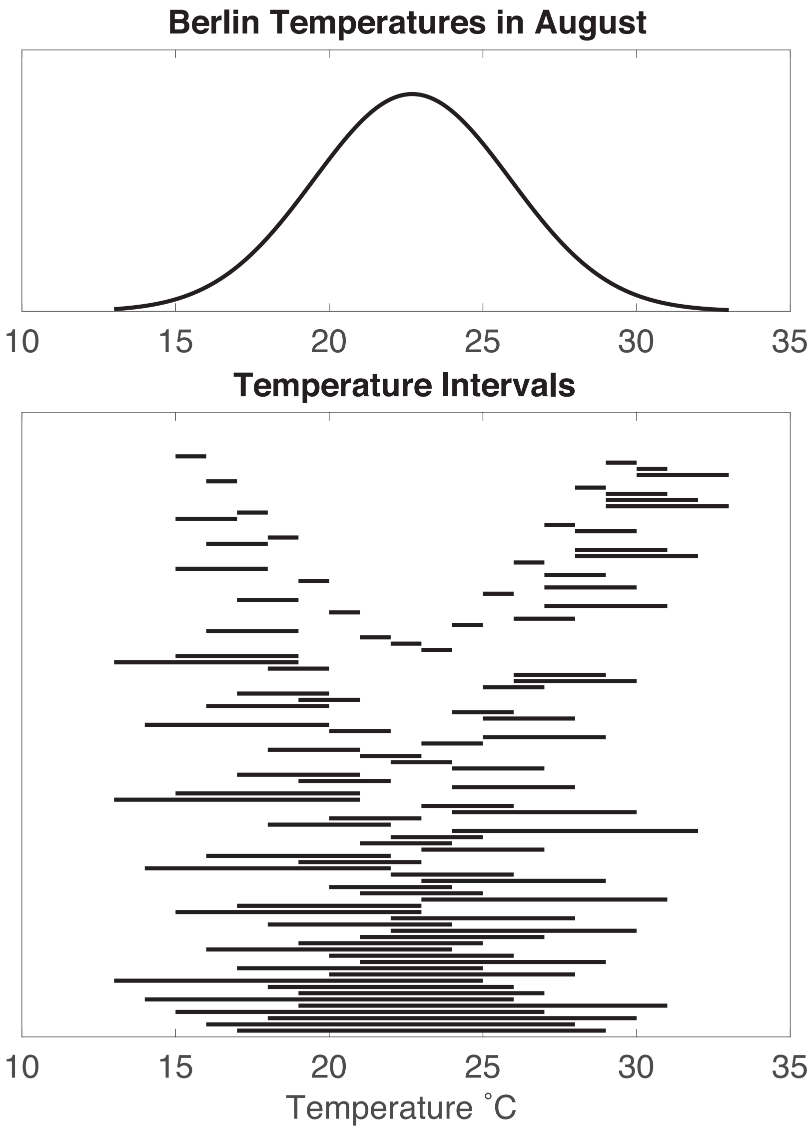


Figure A6. Temperature ranges used in study 3. All bids referred to Berlin weather in August 2015. The top panel shows the distribution of high temperatures in August of 2015. The bottom panel shows the temperature ranges that were used as events in Study 3.

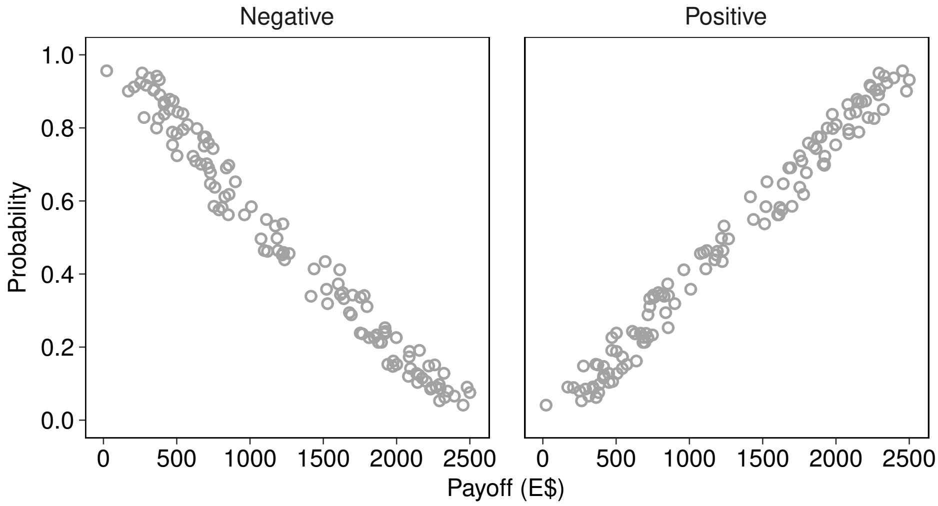


Figure A7. Experiment 3 stimuli used in the exposure phase. Experiment 3 did not have any common gambles included. Probabilities are based on temperature ranges (see previous figure).

Experimental setup

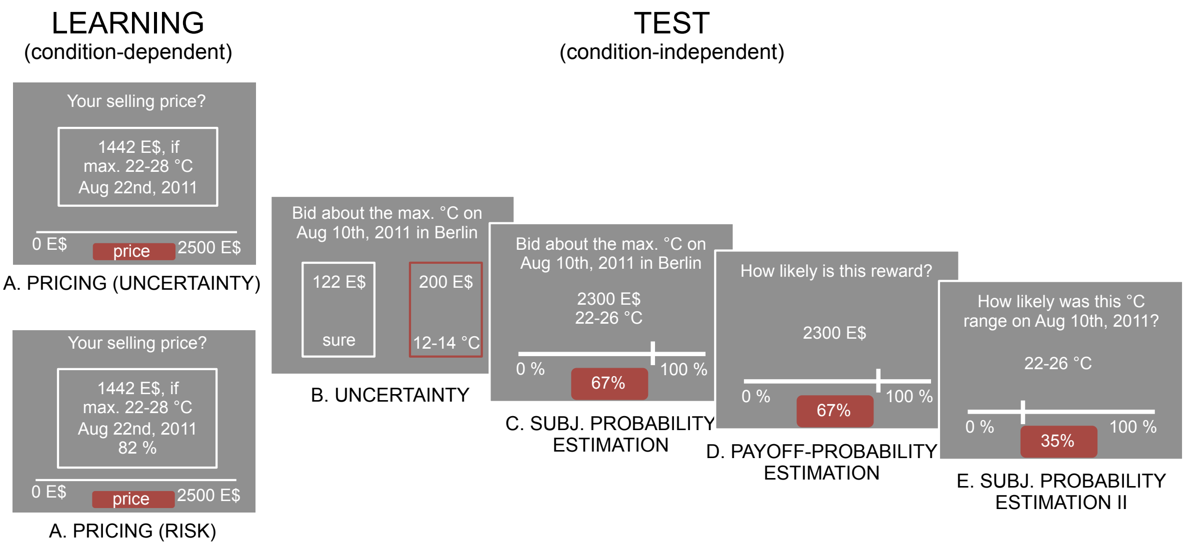


Figure A8. Tasks in Experiment 3. (A) Participants completed a learning phase in which they priced gambles based on Berlin weather drawn from one of two risk–reward conditions. In the pricing under uncertainty conditions, probability information was not given. (B) Then, all participants chose between an uncertain option and an approximately half-as-large certain option pertaining to Berlin weather (15 trials). The task also included a block in which participants chose between an uncertain option and a certain option pertaining to Dushanbe weather (unlearned context, 15 trials). (C) Participants estimated the probabilities for winning the gambles they had also evaluated in the uncertainty task. Again, one block pertained to Berlin weather (15 trials), one block pertained to Dushanbe weather (15 trials). (D) Participants estimated the probabilities associated with 20 different payoffs. (E) Participants estimated the probabilities for historical frequencies they had also evaluated in the uncertainty task (no payoff information). Again, one block pertained to Berlin weather (15 trials), one block pertained to Dushanbe weather (15 trials). The location of payoffs and all other information was randomized between participants (but same location across all tasks). The order of Berlin/Dushanbe blocks was randomized between participants.

Do people transfer learned risk–reward relationships to other contexts?

As mentioned in the main manuscript, we were interested in whether participants would transfer learned risk–reward relationships from the learned context — gambles about Berlin weather — to other contexts. To this end, the tasks in the test phase additionally included blocks that dealt with gambles about Dushanbe weather. We report detailed methods and compare results for both contexts below (by test phase task).

Decisions under uncertainty (methods)

Gambles varied in terms of participants' familiarity with the context. Half of the gambles were about Berlin weather ($N = 15$); the other half were about the weather in Dushanbe, Tadjikistan ($N = 15$). Being based in Berlin at the time of the experiment, participants were more familiar with Berlin weather.This assumption was confirmed in a short post-experiment questionnaire asking participants about their ability to judge Berlin and Dushanbe weather, see section: Manipulation checks. Gambles were presented in blocks (Berlin context & Dushanbe context); context was counterbalanced between participants.

Subjective probability estimation task (methods)

This task consisted of two parts. First, participants were asked to estimate their subjective probability ($0-100\%$) of winning the gamble (i.e., the event occurring) in the decisions under uncertainty task *with payoff information*. Our key interest was the degree to which participants used the payoff information in their estimates. Participants were therefore shown the actual gamble (e.g., "E\$2000 if the maximum temperature in Berlin was between 23 and 26°C on August 22nd") and asked to judge the probability that they would win. Participants completed this task for both the familiar context (Berlin) and the unfamiliar context (Dushanbe) in counterbalanced blocks.

In a second part, participants indicated their subjective probability ($0-100\%$) that the maximum temperature on a given day in August would fall in a given temperature range *without payoff information* (e.g., "likelihood the maximum temperature in Berlin was between 23 and 26°C on August 22nd"). The temperature ranges were identical to those used in the decisions under uncertainty task and in the subjective probability estimation task *with* payoff information. Again, we collected estimates for both the familiar context (Berlin) and the unfamiliar context (Dushanbe) in counterbalanced blocks.

Decisions under uncertainty (results)

Did the experienced risk–reward relationships shape preferences under uncertainty? We expected this to be the case after "learning under risk" (as in Experiments 1 & 2), in particular, but also (though less strongly) after learning under uncertainty. For clarity in reporting and interpreting the results, we ran separate analyses for choices about Berlin (Figure A9A, B) and Dushanbe (C, D). In particular, we analyzed condition-dependent choices after controlling for the events' historical frequencies. Figure A9 shows the results of this analysis. Overall, participants were less likely to choose the gamble over the sure outcome for high (E\$2000) than for low payoffs (E\$100) across conditions ($b_{\text{E\$2000}} = -1.23$, CI =

[−1.62, −.85]). Was this payoff effect moderated by learned risk–reward structures? Indeed, consistent with the risk–reward relationship they had experienced in the learning phase, this payoff effect was smaller for participants who had been exposed to a positive risk–reward relationship under risk (panel A). This effect was driven by participants in the positive condition choosing the gamble less often when the choice was associated with a $E\$100$ payoff—a payoff that had previously been associated with a low probability ($M_{\text{gamble}} = -.12$, $b_{\text{positive} \times \text{E\$100}} = -.57$, CI = [−1.08, −.05], all results based on a mixed effects logistic regression controlling for historical frequencies, using learning type [risk vs. uncertain] × risk–reward relationship [negative vs. positive] × payoff level as predictors).

Learning under uncertainty did not affect choice in either the familiar context (panel B) or in the unfamiliar context of Dushanbe (overlapping credible intervals across panels C, D).

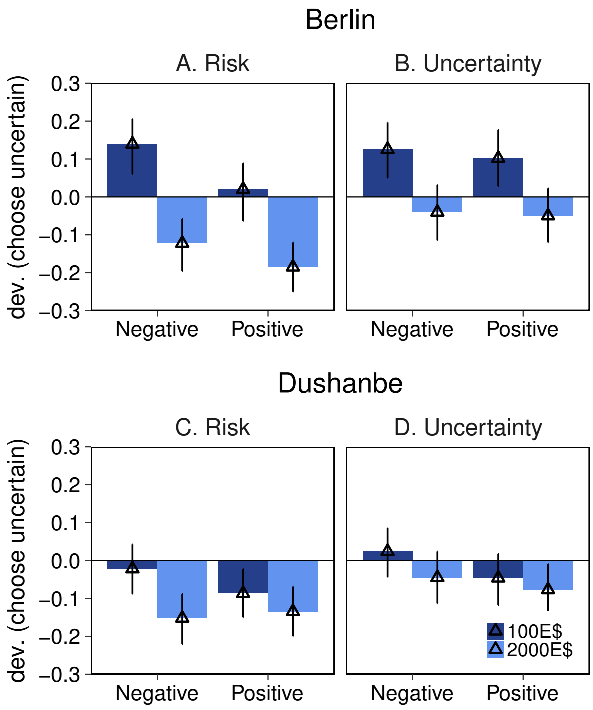


Figure A9. Decisions under uncertainty. Diagrams show how often participants picked the uncertain gamble after controlling for the gamble's historical frequencies (derived from the temperature range). Choice proportions perfectly adjusted to the gambles' historical frequency should have a 0 deviation. Participants in all conditions chose the uncertain option more often when stakes were low (payoff effect $E\$100 > E\$2000$). Bars and triangles reflect the mean of the posterior predictive choice distribution (controlling for historical frequency); error bars indicate the 95% posterior predictive distribution. Posterior predictions were generated using historical frequencies of .5.

Subjective probability estimation tasks (results)

In Experiment 3, participants were asked to estimate the chances that a maximum temperature would fall within a given temperature range both within the context of the gamble as a whole (including payoff information associated with the event) and without this payoff information. As we were interested in how the estimates were affected by the risk–reward environments after controlling for the historical frequencies associated with the events, we report deviations from those historical frequencies. For clarity in reporting and interpreting the results, we ran separate regression analyses for estimates about Berlin (Figure A10

A, B) and Dushanbe temperatures (C, D).

Did participants rely on previously experienced risk–reward structures when gauging their chances of winning a bet about the weather? Figure A10 (A, B) shows that participants' subjective estimates were indeed guided by the payoff information. In line with our predictions and Experiments 1 and 2, in the negative conditions (panel A, left bars), subjective probability estimates were lower when temperature ranges were presented in a gamble context that offered a E$2000 payoff ($b = -.10$, CI = $[-.12, -.07]$) than in a gamble context that offered a E$100 payoff.

This payoff effect—a difference in estimates for $E\$2000$ vs. $E\$100$, after learning under risk—was not observed in the positive condition (panel A, right bars, $b = -.04$, CI = $[-.09, .03]$ modeled in a normal link regression using using learning type [risk vs. uncertain] × risk–reward relationship [negative vs. positive] × payoff level as predictors). However, as Figure A10 shows—and contrary to our predictions—the payoff effect did not flip (with higher payoffs leading to a positive deviation and lower payoffs leading to a negative deviation). A bi-product of this was that participants in the risky positive condition ended up with estimates closer to the true historical frequencies (Figure A10, panel A). We used unstandardized estimates for clarity. Qualitatively, conclusions remained the same when we used logit transforms of the estimates—the model used in Figure A10. For participants who had learned about risk–reward relationships under uncertainty (panel B), the between-condition effects were comparable (larger payoff effect in the negative condition, see panel B). Although a similar pattern of results was observed for the unfamiliar context Dushanbe (C, D), there were no credible interaction effects between payoff and learning phase in this context.

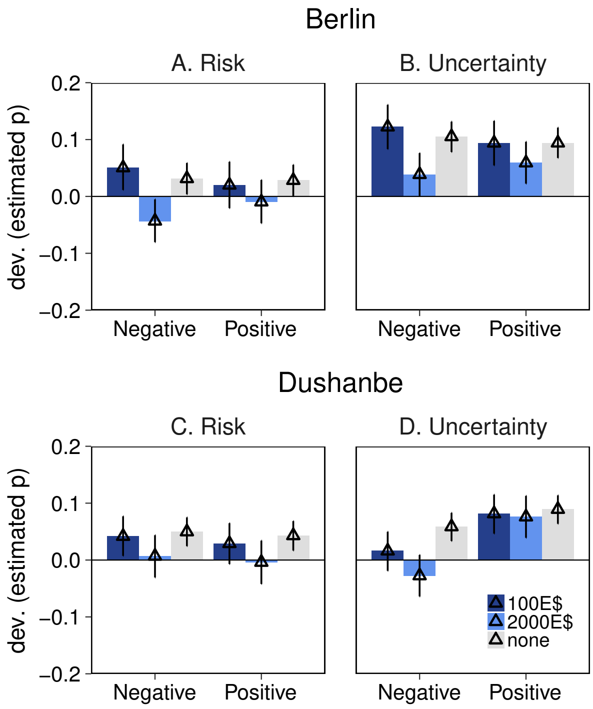


Figure A10. Subjective probability estimation tasks. Plots show deviations of participants' estimates after controlling for the gambles' historical frequencies. Estimates perfectly in line with the gambles' historical frequency should have a 0 deviation. Participants gave subjective probability estimates of winning a particular temperature bet including a payoff (blue bars) or the probability of a given temperature range alone (gray bars). Bars show posterior mean deviations from historical frequencies; error bars show 95% highest density intervals. Posterior predictions were generated using historical frequencies of .5.

Manipulation checks: How well did participants know Dushanbe?

In a short questionnaire after the experiment we asked participants to indicate whether they knew Dushanbe, Tadjikistan (yes/no). This revealed that approximately half of the participants (M = .46, N = 86/186) knew Dushanbe, Tadjikistan. We then asked people to rate their ability to estimate the weather in both Dushanbe and Berlin in 2011. As expected, weather estimation in both contexts was rated difficult, likely because all bets referred to five years in the past (from the time the experiment was run). Specifically, participants' rated ability in estimating Dushanbe weather was Md = 2.00 (M = 3.32, SD = 2.06) on a $1-10$ scale, compared to Md = 3.00 (M = 4.89, SD = 2.25) for Berlin. These results are based on responses of 186/200 participants who filled out the questionnaire.

Decisions under Uncertainty results by historical frequencies

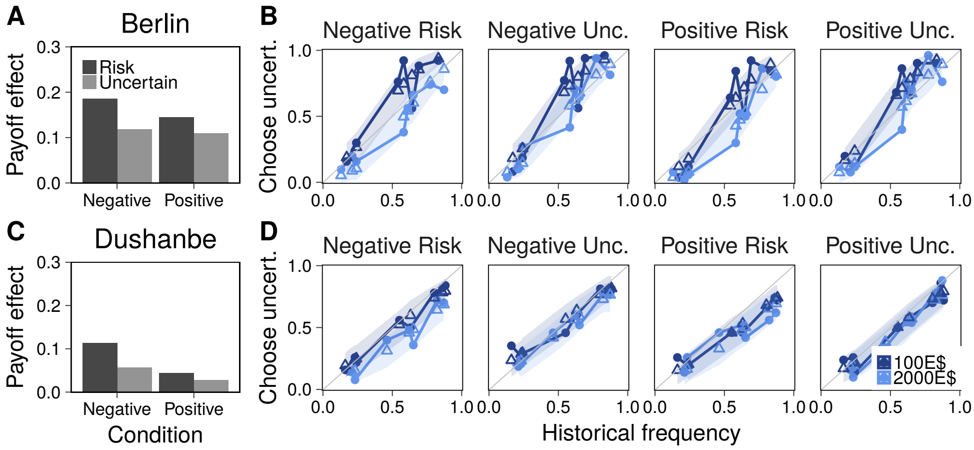


Figure A11. Decisions under uncertainty across different historical frequencies. Participants chose between an uncertain, larger payoff (a weather bet), and a smaller, sure payoff. (A) Payoff effect on choice. Bars depict payoff-dependent differences in gamble choices. On average, participants gambled more for 100*E*$ than for 2000*E*$ bets, but the magnitude on this effect depended on the condition. (B) Payoff effect shown conditional on probabilities (approximated using historical frequencies). (C, D) Results for Dushanbe.

Probability estimation results by historical frequencies

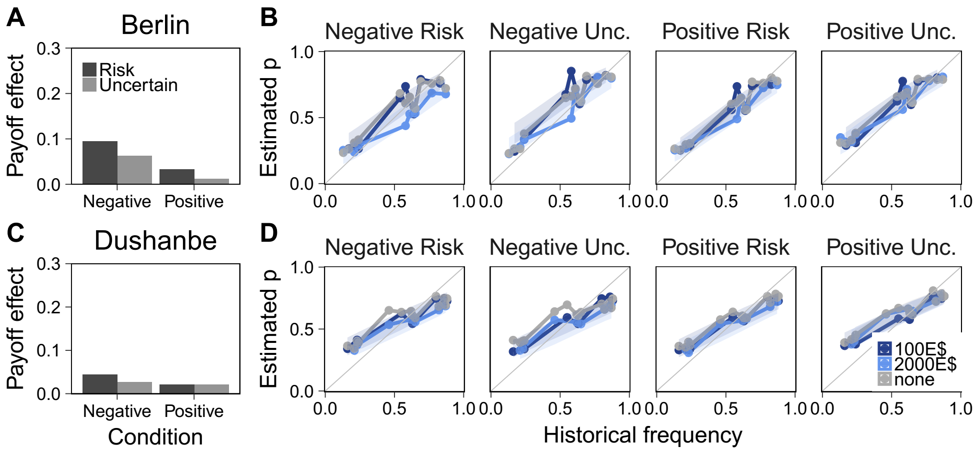


Figure A12. Results of estimation tasks across different historical frequencies. Top row depicts results for Berlin. (A) Average payoff effect on estimates. Bars show aggregated payoff-dependent differences in subjective probability estimates. (B) Payoff effect shown conditional on the probabilities (approximated using historical frequencies). (C, D) Results for Dushanbe.

4. All Experiments

Individual differences in risk–reward estimates and choice

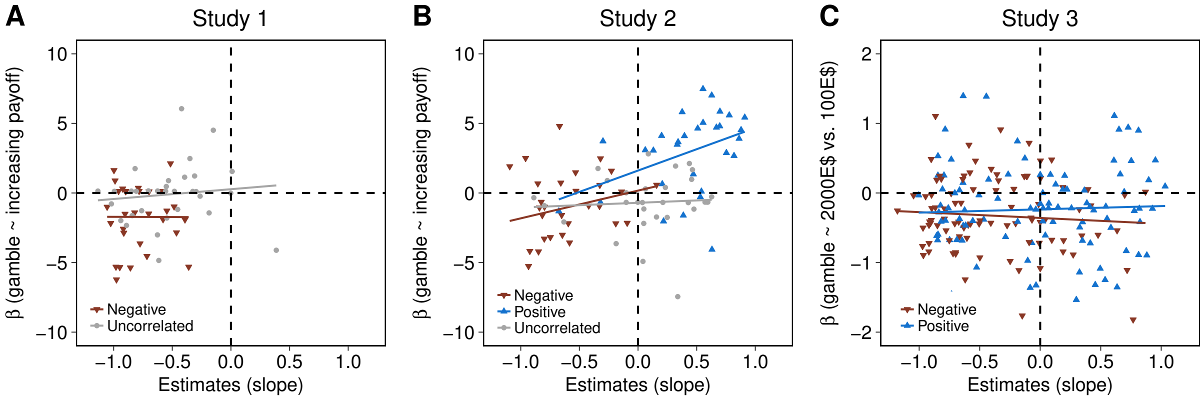


Figure A13. Individual differences in risk-reward estimates and relationship to subsequent choice. (A) Experiment 1. Participants in both conditions estimate mainly negative risk-reward relationships but these are only associated with more conservative choices as payoffs increase in the negative condition. (B) Experiment 2. Slopes predict choices in the positive and negative but not uncorrelated condition. Most slopes reflect the risk-reward condition that participants had been exposed to. (C) Experiment 3. Individual participants showed great variability in risk-reward estimates across conditions. Estimates in the positive condition only weakly predicted choices for the higher payoff gamble, compared to the negative condition (in a model extracting beta's for choices, controlling for empirical probabilities). Differences between conditions not credible. Inferential statistics for all three experiments are reported in the manuscript.

References

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Luce, R. D. (1959). *Individual Choice Behavior*. John Wiley Sons, Inc., New York, NY.

McFadden, D. (1976). Quantal choice analysis: A survey. *Annals of Economic and Social Measurement*, 5:363–390.

Prelec, D. (1998). The Probability Weighting Function. *Econometrica*, 66(3):497–527.

Scheibehenne, B. and Pachur, T. (2015). Using bayesian hierarchical parameter estimation to assess the generalizability of cognitive models of choice. *Psychonomic Bulletin & Review*, 22(2):391–407.

Stewart, N., Chater, N., and Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, 53(1):1–26.

Stewart, N., Reimers, S., and Harris, A. J. (2015). On the origin of utility, weighting, and discounting functions: How they get their shapes and how to change their shapes. *Management Science*, 61(3):687–705.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323.

Yechiam, E., Ashby, N. J. S., and Pachur, T. (2017). Who's biased? A meta–analysis of buyer–seller differences in the pricing of lotteries. *Psychological Bulletin*, 143(5):543–563.

B | Supplementary Material to Chapter 3: “Too good to be true? Psychological responses to surprising options in risk–reward environments”

1. Experiment 1

Pricing (environment gambles)

Condition	**Coefficient (E$) and 95% HDI**
Uncorrelated (baseline)	$-2.27\ (-54.35, 54.26)$
Negative (baseline Unc.)	$29.60\ (-41.34, 102.45)$
Positive (baseline Unc.)	$49.20\ (-22.22, 121.54)$

Table B1. Pricing in Experiment 1 for environment gambles. Coefficients reflect deviations from gambles’ expected values in E$. We accounted for random variation between respondents by including their ID as a grouping factor

Pricing (test gambles)

Main Effects	**Coefficient (E$) and** 95% **HDI**
Surprise vs. Expected	$3.44\ (-24.91; 31.85)$
Surprise vs. Reference	$-11.59\ (-49.15; 26.20)$

Table B2. Pricing in Experiment 1 across gamble types. Coefficients reflect deviations from gambles’ expected values.

Type effects	Type	Coefficient (E$) and 95% HDI
Surprise vs. Expected	Low payoff/low probability	−4.58 (−9.42, 18.55)
	Low payoff/high probability	−19.05 (−34.83, −3.16)
	High payoff/low probability	89.83 (−121.10, 279.14)
	High payoff/high probability	−12.50 (−149.23, 123.27)
Surprise vs. Reference	Low payoff/low probability	−4.91 (−9.06, 18.94)
	Low payoff/high probability	−0.72 (−16.32, 14.87)
	High payoff/low probability	−37.78 (−161.71, 235.56)
	High payoff/high probability	66.40 (−71.46, 204.76)

Table B3. Pricing in Experiment 1 by gamble type. Coefficients reflect deviations from gambles' expected values in E$.

Response Times (test gambles)

Main Effects	Coefficient (s) and 95% HDI
Surprise vs. Expected	3.47 (2.52; 4.43)
Surprise vs. Reference	2.71 (1.64; 3.77)

Table B4. Response times in Experiment 1 across gamble types. Coefficients reflect deviations from individual median RTs in seconds.

Type effects	Type	Coefficient (s) and 95% HDI
Surprise vs. Expected	Low payoff/low probability	3.12 (1.03; 5.21)
	Low payoff/high probability	2.06 (0.09; 4.02)
	High payoff/low probability	4.72 (2.36; 7.08)
	High payoff/high probability	3.93 (1.70; 6.17)
Surprise vs. Reference	Low payoff/low probability	2.45 (0.34; 4.24)
	Low payoff/high probability	2.30 (0.37; 4.43)
	High payoff/low probability	4.02 (1.67; 6.40)
	High payoff/high probability	2.87 (0.63; 5.12)

Table B5. Response times in Experiment 1 by gamble type. Coefficients reflect deviations from individual median RTs in seconds.

2. Experiment 2

Pricing (environment gambles)

Condition	**Coefficient (E$) and 95% HDI**
Uncorrelated (baseline)	76.90 (12.03, 140.71)
Negative (baseline Unc.)	−32.83 (−121.68, 56.75)
Positive (baseline Unc.)	17.29 (−72.91, 108.62)

Table B6. Pricing in Experiment 2 for environment gambles. Coefficients reflect deviations from gambles' expected values in E$.

Pricing (test gambles)

Main Effects	**Coefficient (E$) and 95% HDI**
Surprise vs. Expected	29.05 (0.87; 57.44)
Surprise vs. Reference	22.56 (−25.81; 101.41)

Table B7. Pricing in Experiment 2 across gamble types. Coefficients reflect deviations from gambles' expected values.

Type effects	**Type**	**Coefficient (E$) and 95% HDI**
Surprise vs. Expected	Low payoff/low probability	6.05 (−14.65; 39.35)
	Low payoff/high probability	−0.73 (−19.61; 18.27)
	High payoff/low probability	208.98 (−8.55; 423.59)
	High payoff/high probability	−100.91 (−233.65; 32.16)
Surprise vs. Reference	Low payoff/low probability	1.46 (−18.94; 21.76)
	Low payoff/high probability	−5.22 (−22.53; 14.18)
	High payoff/low probability	95.93 (121.08; 311.87)
	High payoff/high probability	−36.24 (−170.36; 96.89)

Table B8. Pricing in Experiment 2 by gamble type. Coefficients reflect deviations from gambles' expected values.

Response Times (test gambles)

Main Effects	**Coefficient (s) and 95% HDI**
Surprise vs. Expected	0.28 (0.03; 0.53)
Surprise vs. Reference	0.47 (0.08; 0.88)

Table B9. Response times in Experiment 2 across gamble types. Coefficients reflect deviations from individual median RTs in seconds.

Type effects	**Type**	**Coefficient (s) and 95% HDI**
Surprise vs. Expected	Low payoff/low probability	0.50 (−0.19; 1.19)
	Low payoff/high probability	−0.11 (−0.89; 0.69)
	High payoff/low probability	−0.25 (−1.14; 0.64)
	High payoff/high probability	0.94 (0.03; 1.85)
Surprise vs. Reference	Low payoff/low probability	0.32 (−0.37; 0.99)
	Low payoff/high probability	−0.34 (−1.15; 0.48)
	High payoff/low probability	−0.02 (−0.91; 0.87)
	High payoff/high probability	1.06 (0.15; 1.97)

Table B10. Response times in Experiment 2 by gamble type. Coefficients reflect deviations from individual median RTs in seconds.

Pupil Dilation

Pupil dilation (mean vs. peak)

We report changes in median pupil size in the main manuscript. The following two figures show that we obtain qualitatively very similar results when using the mean change (Figure B1), or peak changes (Figure B2).

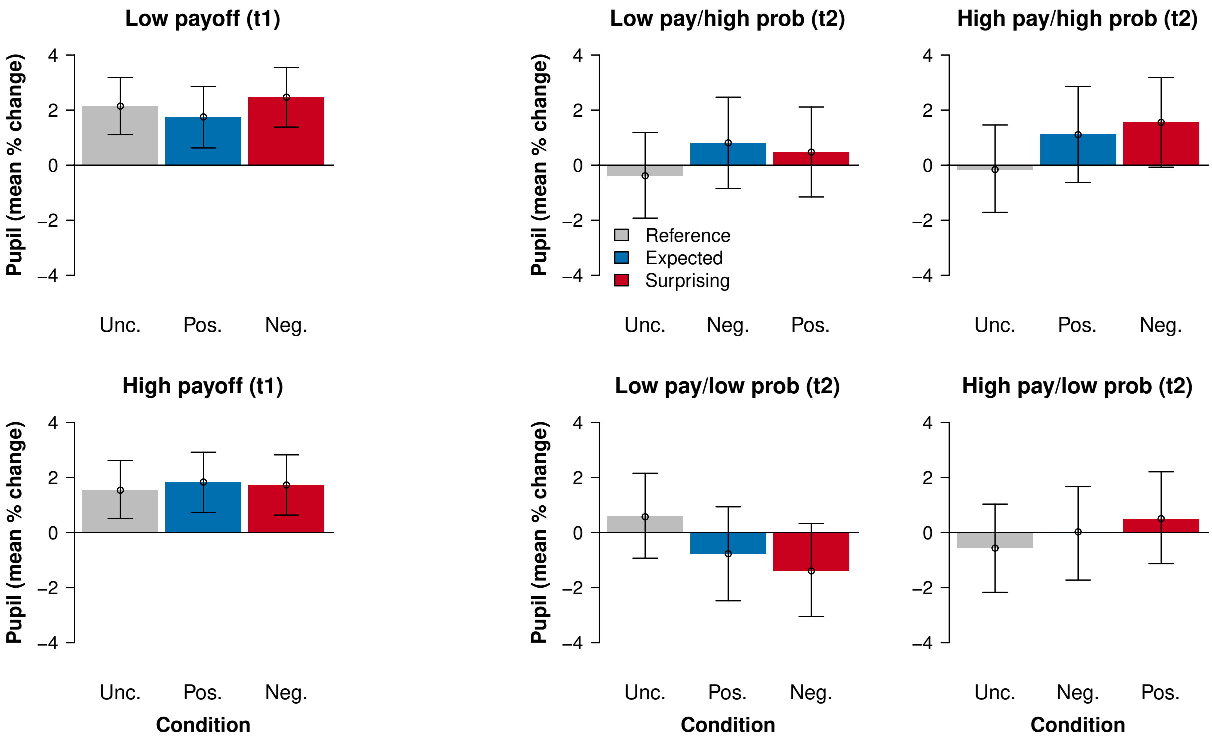


Figure B1. Mean changes in pupil dilation between 1 and 2s after observing payoffs only (left plots, t1), or payoff–probability combinations (right plots, t2). Black triangles and error bars represent the mean and the 95% credible interval of the posterior predictive distribution.

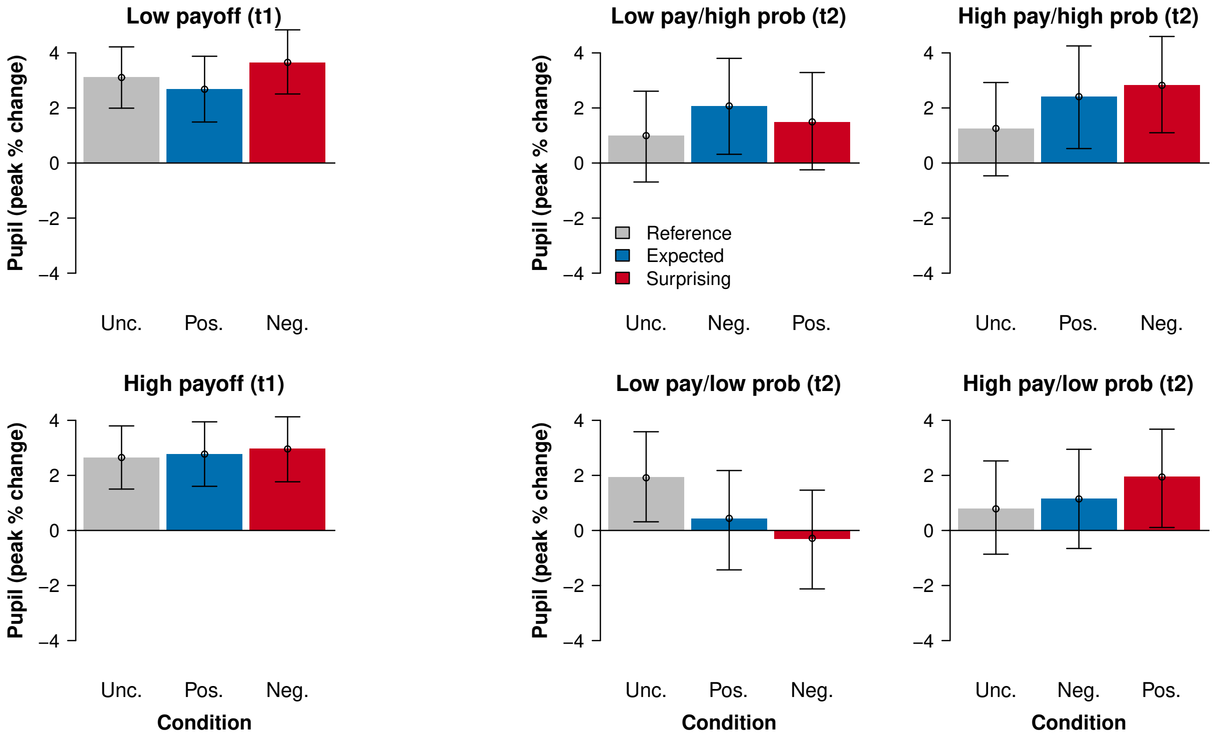


Figure B2. Peak changes in pupil dilation between 1 and 2*s* after observing payoffs only (left plots, t1), or payoff–probability combinations (right plots, t2). Black triangles and error bars represent the mean and the 95% credible interval of the posterior predictive distribution.

Pupil dilation (temporal dynamics)

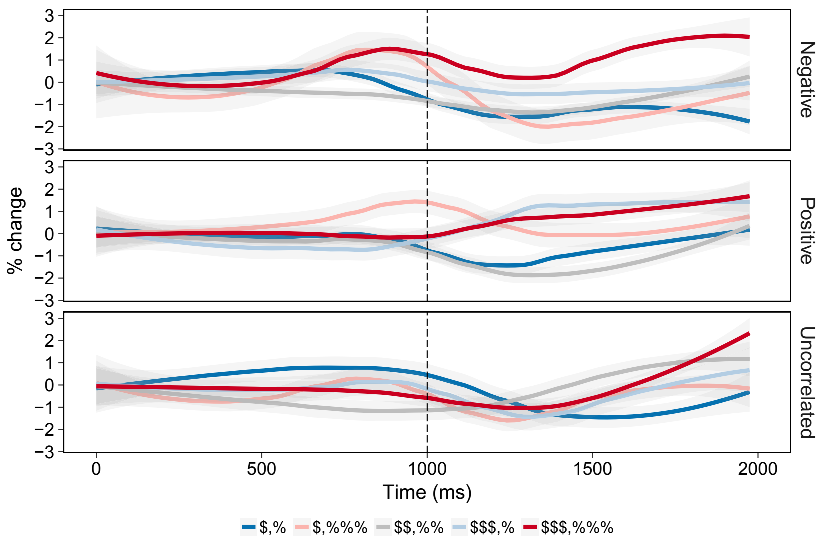


Figure B3. Temporal dynamics of pupil dilation for different payoff–probability combinations across the three conditions (notation $,$$, and $$$ for low, intermediate and high payoffs, %,%%, and %%% for low, intermediate and high probabilities). High payoff/high probability gambles are plotted in red, the low payoff/low probability gambles are plotted in blue. Pupil dilation aggregated across participants and trials and smoothed with a loess function.

Pupil dilation and EV (exploratory)

Is pupil dilation linked to expected value? What is the effect of expected value on pupil dilation across conditions, using all gambles (including environment gambles)? Note that due to the design of the experiment, the (marginal) distribution of expected values differs across the three conditions. Specifically, in the negative condition, expected values never exceeded E$600, except for the high payoff/high probability gambles, with expected values close to E2500. These gambles appeared after 40 environment-only trials. Therefore, high payoff/high probability gambles may have elicited EV surprise in the negative condition. Indeed pupil dilation was linked to EV in the negative condition (b = .0009, CI = [.0003, .0014]). The marginal distribution of EVs in the positive condition was also bimodal. Again, expected value was linked to pupil dilation (b = .0006, CI = [.0003, .009]). For both of these correlated conditions, the effect was small. Expected value was not linked to pupil dilation in the uncorrelated condition (b = .000, CI = [−.0004, .0003]).

C | Supplementary Material to Chapter 4: "Risk–reward structures can promote satisficing in decisions under risk"

1. Incidental learning

Incidental learning task (setup)

Participants indicated their willingness to sell (WTS) for one gamble at a time, with self-paced breaks between five blocks. To motivate participants to indicate their true valuations of the gambles, we implemented a Becker-DeGroot-Marschak auction (Becker et al., 1964): Participants entered a price at which they would be willing to sell each gamble by moving the mouse along a rating scale ($0 - 100E\$$) and confirming the value with a click. 10 gambles were played out the end of the experiment. For those 10 gambles, the experimenter then offered a randomly generated buying price between 0 and the absolute payoff in that gamble. If the experimenter's buying price exceeded the participant's selling price, participants sold the gamble and earned the buying price. Otherwise, the gamble was played out (e.g., 50% chance of $38E\$$). The dominant strategy in this task is to price a gamble based on its subjective value: Setting higher prices can prevent selling unattractive gambles; lower prices can lead to selling attractive gambles under value. Experiments coded in PsychoPy (Peirce, 2007).

Posttests (setup)

In the posttest phase, as an independent test of how well (individual) participants had picked up and maintained the different risk-reward relationships, participants completed three tasks. First, we asked participants to choose between an uncertain option and a sure thing (sure payoff = .3 × uncertain payoff). Second, we asked participants to estimate probabilities from payoffs. Specifically, we asked them to estimate the likelihood they would win 20 randomly drawn payoffs that fell within the range of possible payoffs within the experiment. Third, we reversed this task and asked them to estimate the payoffs that had been associated with 20 different probabilities (as in Leuker et al., 2018a). We use the probability estimation task to estimate how well people learned (and maintained) the condition–dependent risk–reward structures over the experiment (see statistical analyses below).

Posttests (results)

Figure C1 plots how often people chose the uncertain option, participants' payoff–dependent probability estimates and their probability–dependent payoff estimates. Choices were impacted by the previously experienced risk–reward structure. In comparison to our earlier study (Leuker et al., 2018a), the uncertain option was chosen more often overall since in this experiment, the sure thing was made less attractive, by offering only 1/3 of the payoff that the uncertain option offered; the ratio was 1/2 in our earlier studies.

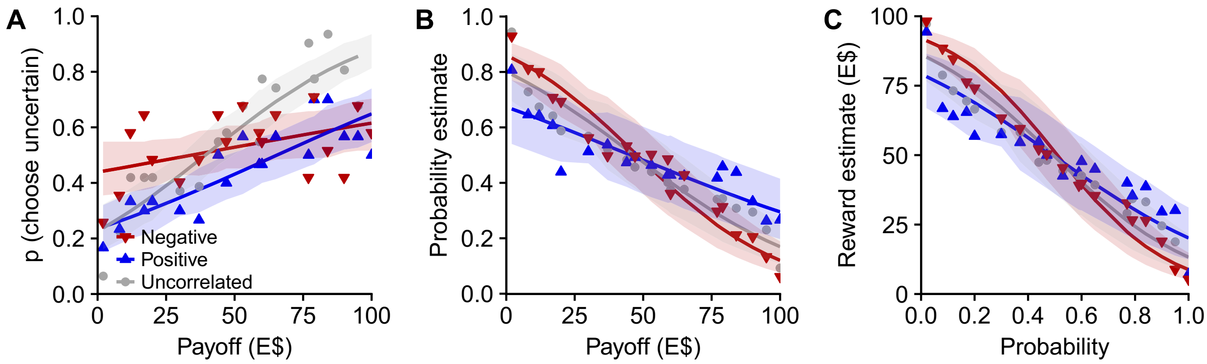


Figure C1. Posttests. (A) Participants decisions under uncertainty were impacted by the risk–reward structures they had been exposed to previosly. (B, C) Payoff and probability estimates were influenced by the risk–reward structure from the incidental learning phase, but strongly biased towards an inverse relationship between risks and rewards.

The table shows the coefficients of the probability estimates in panel (C).

Condition	Slope (β)	Highest Density Interval (β)
Negative, Incidental	−0.65	(−0.69; −0.61)
Positive, Incidental	−0.28	(−0.34; −0.22)
Uncorrelated, Incidental	−0.58	(−0.62; −0.53)

Table C1. Risk–reward conditions reliably impacted participants estimates ($b_{uncorrelated}$ = .16, CI = [.09, .23], $b_{positive}$ = .37, CI = [.30, .44], model that compares estimates to those in the negative condition), but all estimates were biased to a negative risk–reward structure.

Test phase (gamble pair characteristics)

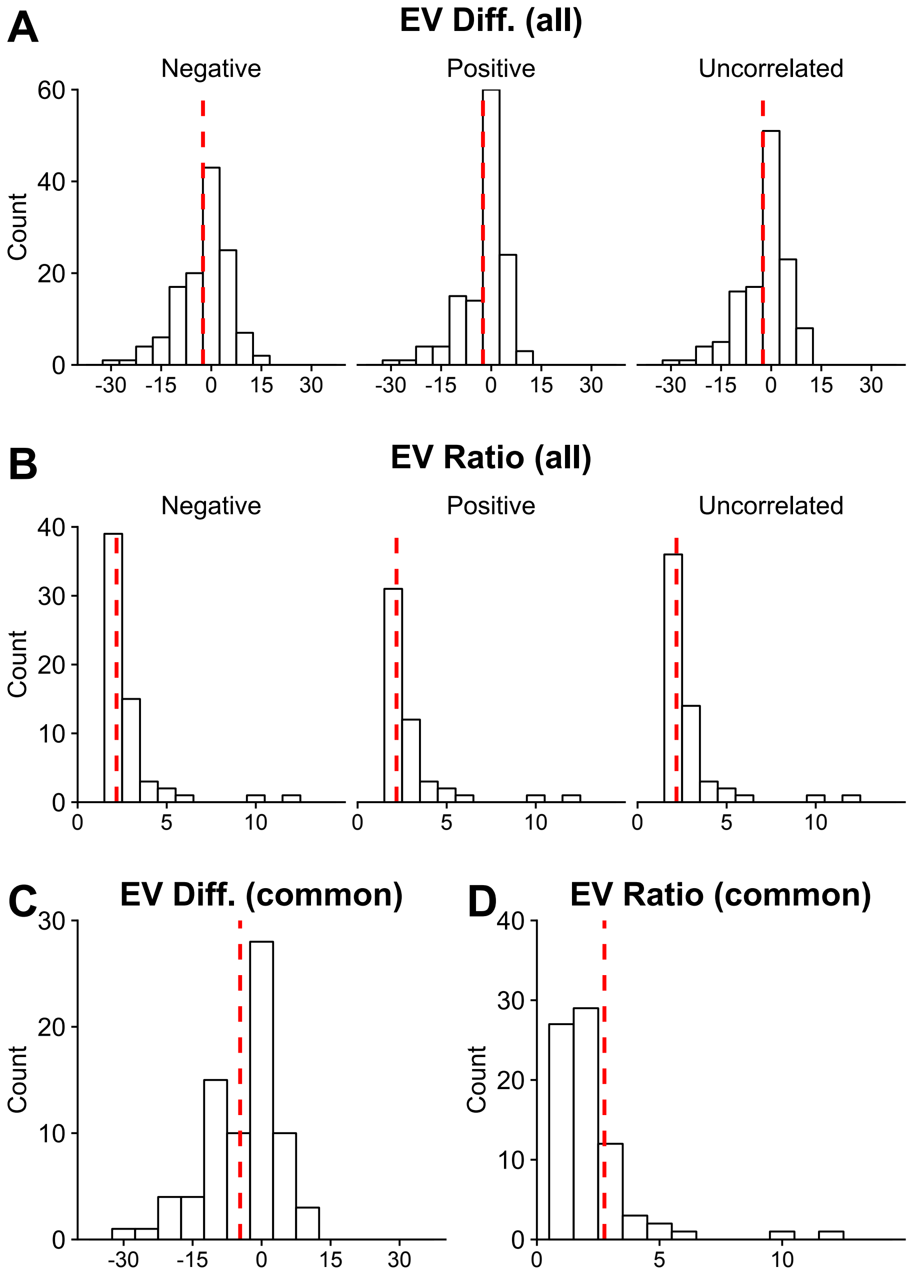


Figure C2. Characteristics of gamble pairs in the choice phase. (A, B) Some differences emerge due to the proportion of gambles that were condition–dependent—such as a larger proportion of gambles with very small EV differences in the positive condition. (C, D) Analyses were based on common gamble pairs.

Test phase (model–choice data link)

Beyond DICs and posterior predictive checks, one way to assess how well the model describes the data is to link individual differences in parameter estimates back to individual differences in the proportion of EV choices. Figure C3 shows that there were plausible links between individual parameter estimates and individual differences in the proportion of times the higher expected–value option was chosen.

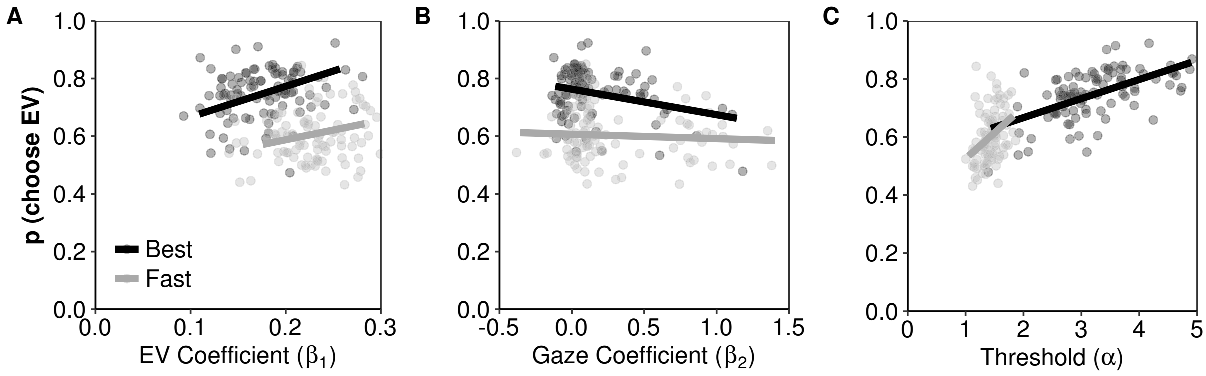


Figure C3. Relationship between DDM parameters and behavioral choice results. Each dot represents one participant. (A) Participants who were sensitive to EV differences chose the higher EV option more often. (B) Participants who distributed their attention more evenly (indicated by a gaze coefficient of 0) chose the higher EV option more often. (C) Participants who set higher thresholds chose the higher EV option more often.

2. Explicit learning

Explicit learning task (hypotheses)

We assessed to what extent our results extended to another learning condition. As mentioned before, data comes from the same experiment, namely from $N = 92$ participants who learned about one of three risk–reward structures explicitly (i.e. using a function learning task).

Sometimes people are aware that their primary task is to learn and have the chance to receive explicit feedback about the underlying structure—for example when they learn about the slim chances of winning (or have to compute them) in school, or even when a friend, colleague or your doctoral supervisor explicitly tells you about the probability of getting a paper accepted in the journal you desire. In laboratory studies on cue–criterion learning (sometimes "intentional learning", Wattenmaker, 1991; Whittlesea, 1987), it has been shown that positive relationships are easier to learn than negative ones (Klayman, 1988). One of our predictions for the learning phase task was that introducing the context of the relationship they are learning (i.e. a risk–reward relationship), participants may learn an inverse risk–reward relationship sooner than a positive one.

Explicit learning task (setup)

Here, such a task was implemented as follows: Participants saw one payoff at a time. Their task was to guess the probability associated with the payoff of the gamble (the probability was covered with a "?", Figure C4). Participants were informed that all gambles are drawn from the same set of gambles. In each round, participants entered their estimates with a mouse click on a rating scale ($0 - 100E\$$) and confirmed them with a click on the value. After each guess, participants received feedback about how close their estimate was to the actual value, but not in which direction they deviated. Participants received points for closer estimates, and could earn 200 points for learning the association (i.e. reaching the criterion $+/-8$) in fewer than 30 trials. After five correct guesses (or 100 trials), participants moved on to the choice task. Participants received points for closer estimates (10 points for a deviation of 0, 9 points for a deviation of 1(...) and no points for deviating more than 8). Correct guesses in the uncorrelated condition were 50% +/− 8). Points translated to bonuses in *E*$ (i.e. 200 points were equal to a bonus of 200*E*$, or €2; bonus rules and exchange rate revealed in instructions). After 100 trials, participants moved on to the next part of the experiment irrespective of having reached the criterion or not. This was the case for almost all participants in the uncorrelated condition.

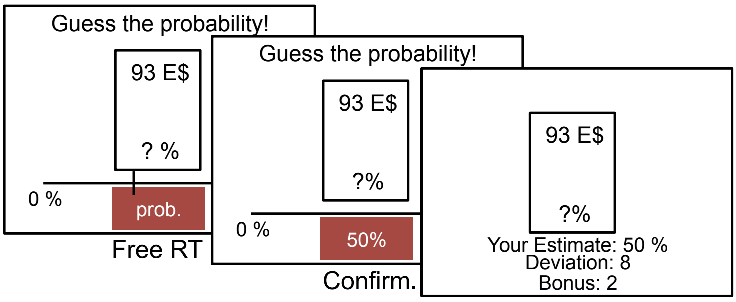


Figure C4. Explicit learning task.

Explicit learning task (results)

In contrast to the idea that priors from nonlaboratory environments would aid the learning of an inverse relationship between risks and rewards, the positive risk–reward relationship was learned faster than the negative risk–reward relationship ($M_{pos.} = 16.8$ trials, $M_{neg.} = 30.13$ trials, difference positive vs. negative: $b = -13.34$, CI $= [-24.37; -2.42]$). As expected, most participants in the uncorrelated risk–reward condition completed all 100 trials, a few participants hit the criterion earlier by indicating 50% on five consecutive trials (resulting in $M = 86.2$ trials). As shown below, the probability estimation task at the end of the experiment revealed that participants' probability estimates reflected the risk–reward structure they had been exposed to previously.

Posttests (results)

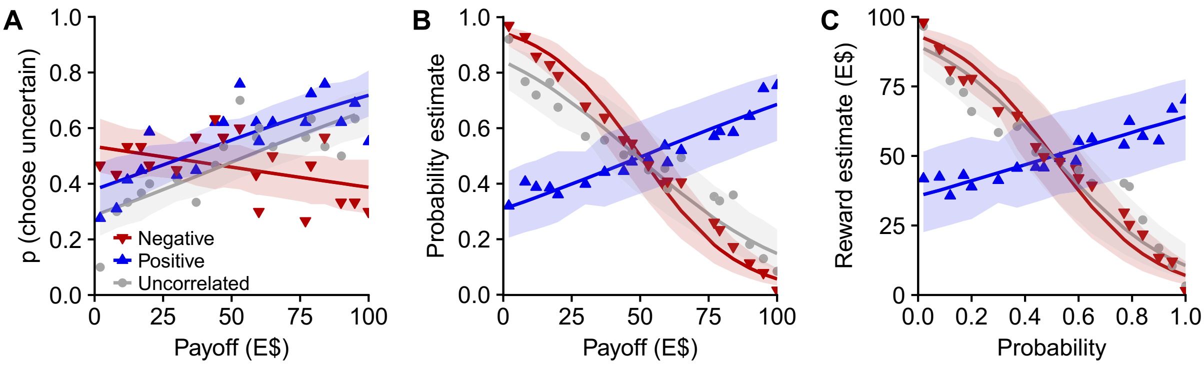


Figure C5. Posttests. (A) Participants decisions under uncertainty were impacted by the risk–reward structures they had been exposed to previosly. (B, C) Payoff and probability estimates were influenced by the risk–reward structure from the incidental learning phase, but in the uncorrelated condition biased towards an inverse relationship between risks and rewards.

The table shows the coefficients of the probability estimates in panel (C).

Condition	**Slope (β)**	**Highest Density Interval (β)**
Negative, Explicit	-0.90	$(-0.93; -0.87)$
Positive, Explicit	0.22	$(0.15; 0.29)$
Uncorrelated, Explicit	-0.49	$(-0.54; -0.44)$

Table C2. Participants' probability estimates reflected the risk-reward structure they had been exposed to previously. The negative condition provided *lower* probability estimates for gambles with higher payoffs and the positive condition provided *higher* estimates for higher payoffs. The uncorrelated condition provided lower estimates for higher payoffs (weaker slope compared to the negative condition; $b_{uncorrelated} = .32$, CI = $[.25, .29]$, $b_{positive} = 1.11$, CI = $[1.04, 1.19]$, model predicting estimates from reward $\times$ Condition interaction, with negative condition as a baseline).

Test phase (descriptive results)

For the test phase, we conducted the same set of analyses as for the incidental learning task. Again, participants in the negative risk–reward condition maximized expected values less than participants in the other two conditions, when making the best choice was emphasized ($b_{unc.>neg.} = 0.24$, CI $= [0.01, 0.47]$; $b_{pos.>neg.} = 0.40$, CI $= [0.17, 0.63]$). There were no reliable response time differences across risk–reward conditions in the best instruction (all CIs included 0), but the positive condition spent slightly more time choosing in the fast instruction ($b_{pos.>neg.} = 0.10$, CI $= [0.03, 0.19]$). Since we did not find robust differences in processing strategies in the fast condition otherwise, we do not interpret this result further.

Again, participants in the negative condition inspected fewer attributes than the other two conditions ($M_{neg.} = 2.77$, CI $= [2.55, 3.03]$). Participants in the other two conditions seemed to sample information more carefully ($M_{pos.} = 3.27$, $b_{pos.>neg.} = 0.50$, CI $= [0.17, 0.84]$); $M_{unc.} = 3.09$, $b_{unc.>neg.} = 0.31$, CI $= [-0.02, 0.65]$). Inspecting more attributes was linked to choosing the higher–EV option across conditions ($b_{best} = 0.13$, CI $= [0.06, 0.19]$; $b_{fast} = 0.06$, CI $= [0.00, 0.13]$, all after controlling for individual variation and expected value differences).

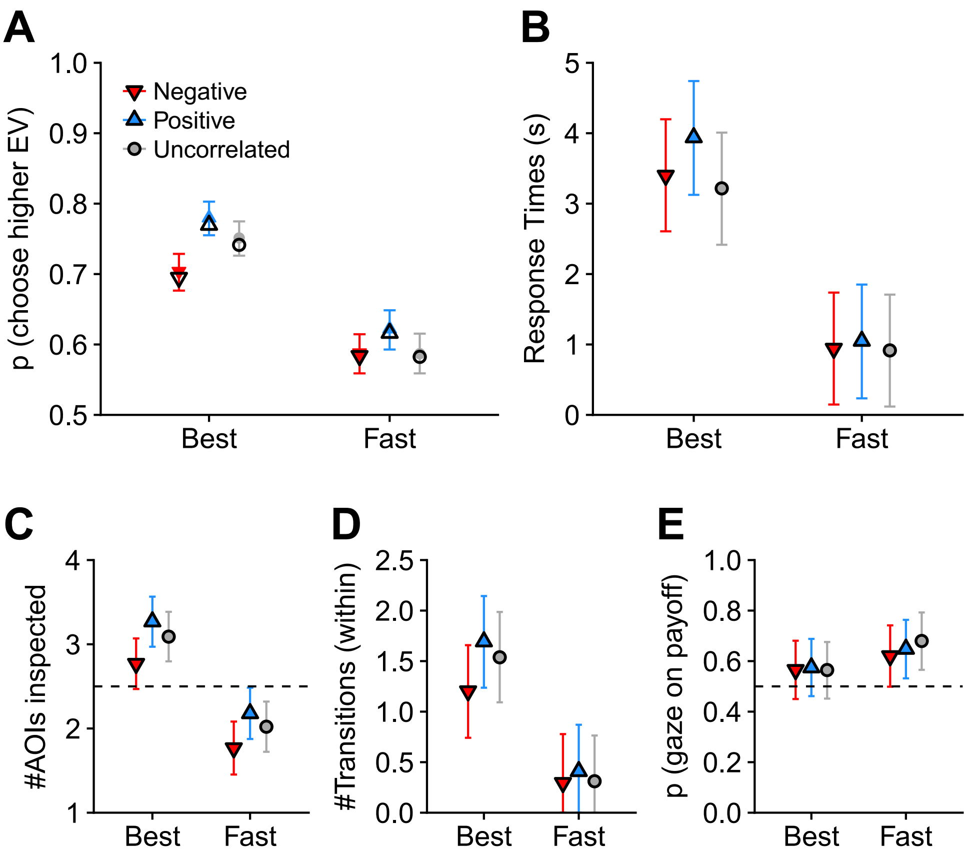


Figure C6. Descriptive results for the test phase after explicit learning.

Test phase (computational model)

After explicit learning, the estimated parameters across the three risk–reward environments are largely comparable across conditions (most CI's include 0, Figure C7). This means that the learned risk–reward environment impacted subsequent learning not the same way incidental learning did: Specifically, participants in the negative risk–reward did not lower their threshold α—i.e. they did not take less time than participants in the other two conditions. Instead, thresholds (Figure C7A) are highest in the positive condition. Moreover, in contrast to the incidental learning conditions, the distribution of gaze impacted evidence accumulation in all conditions (all coefficients > 0), and similarly across conditions (all condition–dependent CI's included 0).

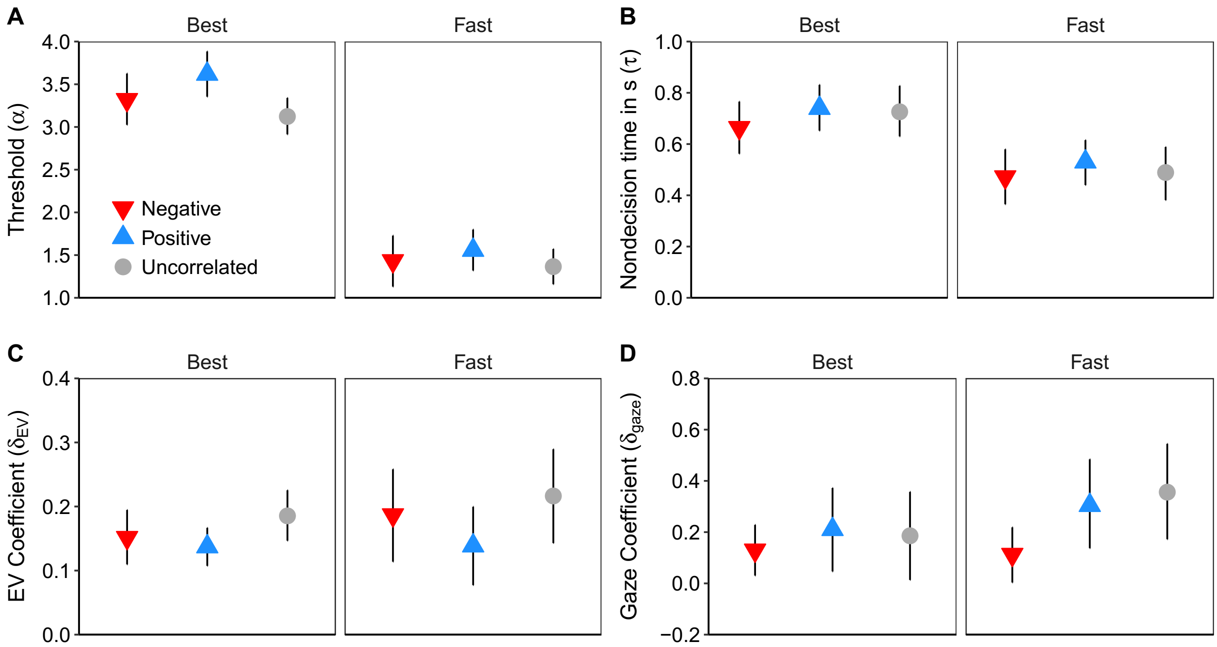


Figure C7. Posterior distributions for the group-level parameter estimates.

Lastly, Figure C8 shows that generally, the parameter estimates from the extended drift diffusion model were consistent with choice patterns of the behavioral data.

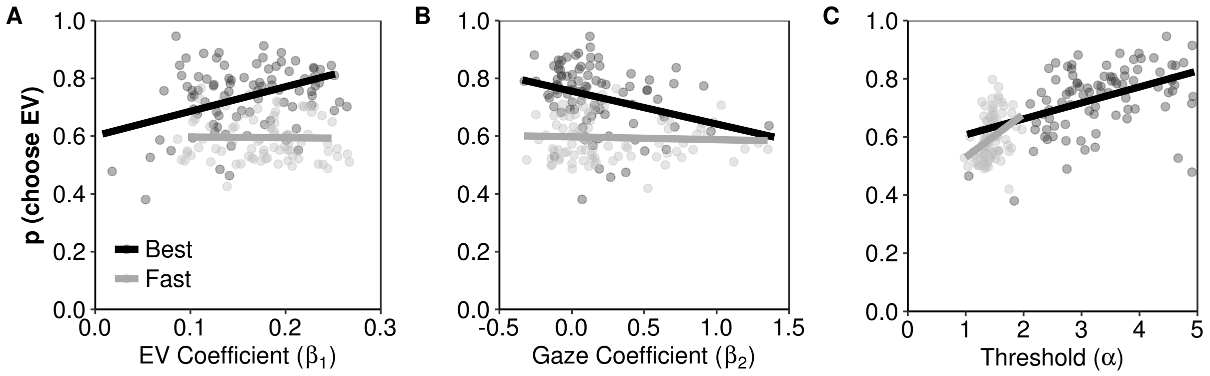


Figure C8. Relationship between DDM parameters and behavioral choice results. Each dot represents one participant. (A) Participants who were sensitive to EV differences chose the higher EV option more often. (B) Participants who distributed their attention more evenly (gaze coefficient of 0) chose the higher EV option more often. (C) Participants who set higher thresholds chose the higher EV option more often.

Test phase (discussion)

What is the conceptual difference between explicit and incidental learning? One way to understand these results is to think back of what participants could learn in incidental learning conditions when pricing gambles from different risk–reward environments. They learn two elements of risk–reward environments: First, they can learn about the functional form of the risk–reward structure (are risks and rewards positively related, negatively related or uncorrelated?). Second, they can also learn about the expected value distributions in a given environment, with a negative risk–reward structure being composed of many options with similar values (\$20 with $p = .8$, and \$80 with $p = .2$). This resembles the structure of the environment outside the lab, in which EVs across options are typically identical in monetary domains with a pay–to–play structure. For instance, the probability of obtaining a reward when there is a \$1 pay–to–play fee is given by $p = 1/(1 + gain)$. This mechanism is for example found in roulette, and gives rise to risks and rewards being inversely related through a power law. In the current experiments, we exposed people to a linear relationship between risks and rewards and thus intermediate values ($50E\$$) had slightly higher EVs than high and low values.

While the functional form can be learned equally well—if not better—in an explicit learning task, there is an additional step involved in then inferring what that means for EV distributions in a given choice environment. It may be the perception of similar EV distributions that led participants in the negative incidental condition to set lower thresholds and sample less information. This is consistent with other work (Leuker et al., 2018b) which indicates that negative risk–reward environments can elicit "EV surprise" for oddballs with higher expected values than usually experienced.

3. Drift diffusion modeling: Model comparisons

General approach

We used a Bayesian Hierarchical Modeling approach to estimate individual and group parameters simultaneously. We used JAGS in R to sample from the posterior distributions. In Bayesian parameter estimation, parameter estimates are represented as prior distributions and then updated into posterior distributions based on the observed data. The advantage of a hierarchical approach is that it can account for individual variation while simultaneously pooling individual estimates into group-level distributions. The joint posterior parameter distributions were estimated using Monte Carlo Markov Chain methods implemented in JAGS, called from R. We ran 25 chains, each with 4,000 recorded samples, which were drawn from the posterior distributions after a burn-in period of 500 samples. Model selection was based on deviance information criteria, with lower values indicating better fit; and on posterior predictive checks in which we assessed how well the model aligned to behavioral results (across learning conditions).

The experiment had a risk–reward (between-participants) × timepressure (within-participants) design that we fit simultaneously. Moreover, half of the participants completed the choice task after incidental learning and the other half after explicit learning. We fit the two learning conditions separately.

We inspected the quality of the posterior distributions by visually inspecting the mixing of the chains and autocorrelation, and on the basis of the Gelman-Rubin statistic. We compared model fits using DIC values (lower values indicate better model fit) and creating posterior predictions for choices and response times using the mean parameter values obtained from the "winning" models, on both the condition and the participant–level.

We compared five different models with the simple DDM as the baseline model. The second model only takes into account the value differences between the options, with larger EV differences leading to higher drift rates. The third model predicts drift from gaze differences between the two options alone. The fourth model tests for an additional main effect of gaze on the drift rate. The fifth model also allows for an interaction between gaze and value (similar to Cavanagh et al., 2014), including additive effects of value and gaze. Model selection was done by fitting the models across learning conditions (to obtain one DIC per model). Parameter estimates were obtained by re-fitting the models separately for each learning condition. We did not include an aDDM–like model in which the drift rate *solely* depends on the interaction between value and gaze, without additive effects—the reason for this is that we wanted to partial out and compare "EV sensitivity"/"EV usage" for the three risk–reward conditions, which is not possible in an interaction–only model. The best–fitting model was an extended DDM in which the drift rate depended on an additive effect of gaze differences and EV differences (i.e. each of which was quantified by a free parameter, or regression coefficient).

Drift diffusion model specifications

Model 1: Simple DDM We estimated the main diffusion parameters for each condition, α, τ, and δ (β is fixed at .5).

Model 2: Value differences impact drift rate

We estimated α and τ, and fix β at .5. In addition, we estimated the drift rate using a regression approach, such that the drift is described by an intercept $\delta 0$ and a regression coefficient, δ_{EV}. The intercept is participant–dependent and therefore represents individual differences in the ability to detect the higher EV option.

$$\delta = \delta_0 + \delta_{EV} \times (EV_H - EV_L) \tag{1}$$

Model 3: Gaze differences impact drift rate

We estimated α and τ, and fix β at .5. In addition, we estimated the drift rate using a regression approach, such that the drift is described by an intercept δ_0 and a regression coefficient, δ_{gaze}. The gaze coefficient models to what extent choices depend on pure gaze towards a particular option. Gaze is entered as the proportion of total gaze to the higher vs. lower EV gamble.

$$\delta = \delta_0 + \delta_{gaze} \times (gaze_H - gaze_L) \tag{2}$$

Model 4: Value and gaze differences impact drift rate

We estimated α and τ, and fix β at .5. In addition, we estimated the drift rate using a regression approach, such that the drift is described by an intercept δ_0 and two regression coefficients, δ_{EV} and δ_{gaze}. The value coefficient models peoples' sensitivity to differences in expected values. The gaze coefficient models to what extent choices depend on pure gaze towards a particular option. Gaze is entered as the proportion of total gaze to the higher vs. lower EV gamble.

$$\delta = \delta_0 + \delta_{EV} \times (EV_H - EV_L) + \delta_{gaze} \times (gaze_H - gaze_L) \tag{3}$$

Model 5: Value, gaze and their interaction impact drift rate (aDDM)

As model 3, but with an interaction between value and gaze ($\delta_{interaction}$). Cavanagh et al. (2014) modeled the interaction without separate additive effects, but here we aimed to compare "EV use" across risk–reward conditions.

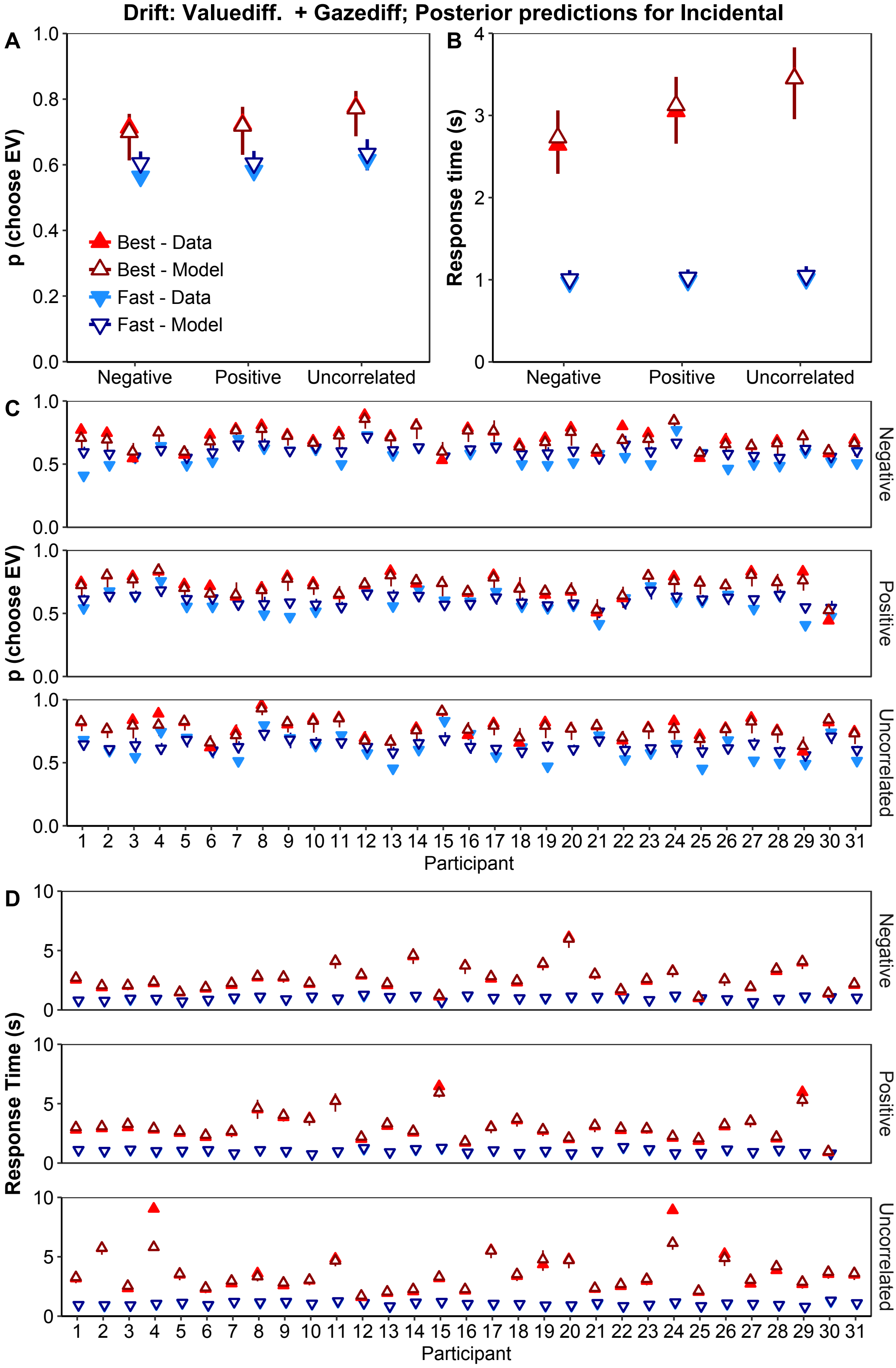
$$\delta = \delta_0 + \delta_{EV} \times (EV_H - EV_L) + \delta_{gaze} \times (gaze_H - gaze_L) +$$

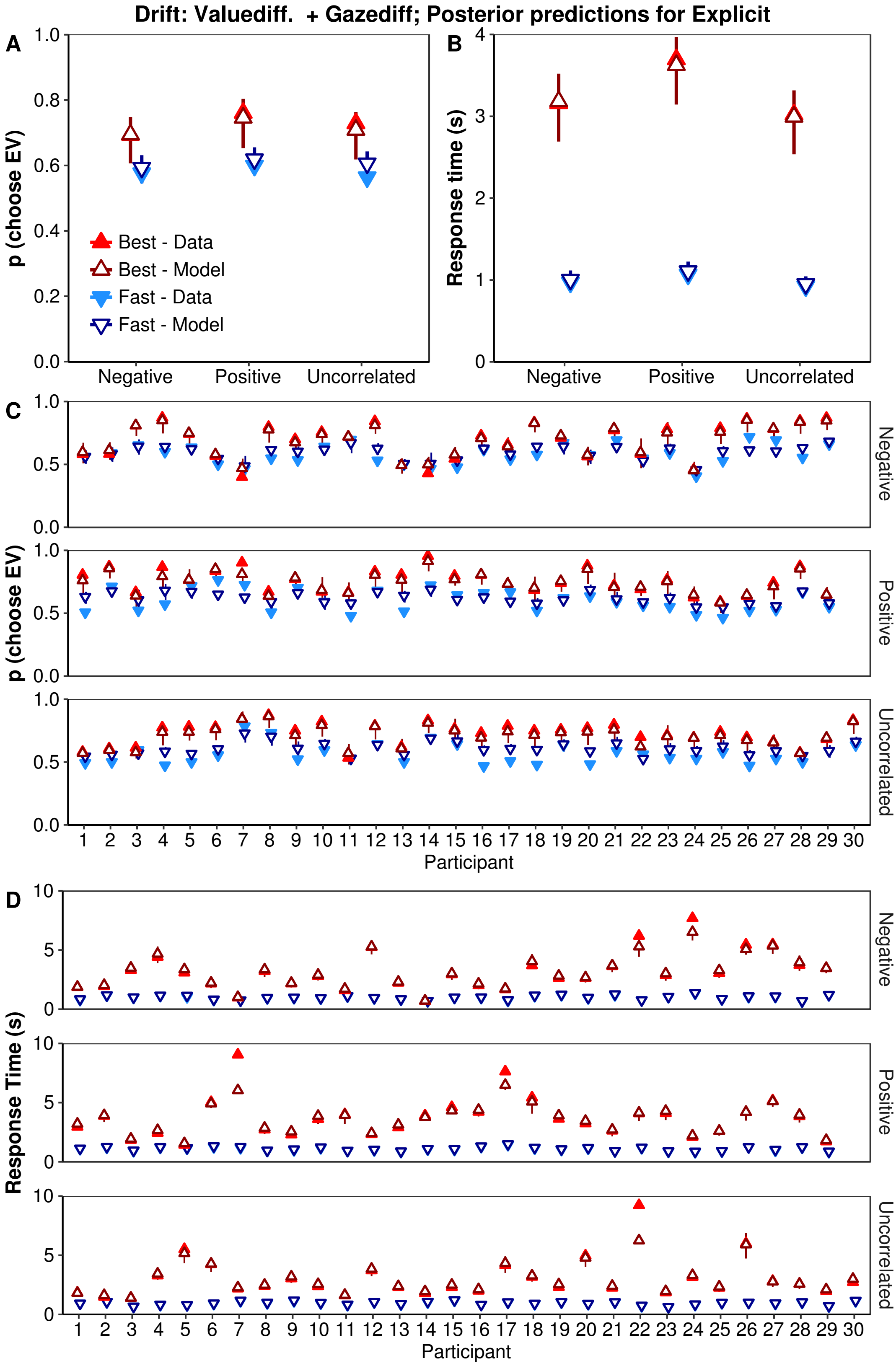
$$\delta_{interaction} \times (gaze_H \times EV_H - gaze_L \times EV_L)(4)$$

Results: DICs and posterior predictions

Model	DIC
Model 1: Simple DDM	73124.41
Model 2: Drift: Valuediff.	72027.83
Model 3: Drift: Gazediff.	72138.86
Model 4: Drift: Valuediff. + Gazediff.	71274.90
Model 5: Drift: Valuediff. + Gazediff. + Value × Gaze	71385.54

Table C3. Deviance Information Criterion (DIC) for five different formalizations of the Drift Diffusion model.




Drift: Valuediff. + Gazediff; Posterior predictions for Explicit
A
B
C
D
p (choose EV)
Response time (s)
Response Time (s)
Best - Data
Best - Model
Fast - Data
Fast - Model
Negative
Positive
Uncorrelated
Participant

4. Alternative modeling account: Prospect theory

Methods

Prospect Theory makes no predictions about response times, but has been fruitful in studying decisions under timepressure or cognitive load previously (Young et al., 2012; Olschewski et al., 2018). As before, a Bayesian Hierarchical approach allowed us to fit the model at the participant level and at the group level simultaneously. We used the a standard power function for the value function, a one-parameter probability weighting function (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), and a Luce choice rule as an error theory (Luce, 1959; McFadden et al., 1999).

Results

Consistent with results from the Drift Diffusion Model (in which participants in the negative risk–reward environment set slightly lower thresholds), participants in the negative risk–reward environment made less deterministic choices (as indicates by the choice rule parameter, θ). There were no substantial systematic differences in peoples subjective evaluations of payoffs and probabilities as a function of risk–reward structures. The result that timepressure leads to less deterministic choices is in line with prior research, finding that cognitive load results in lowered choice consistency (Olschewski et al., 2018).

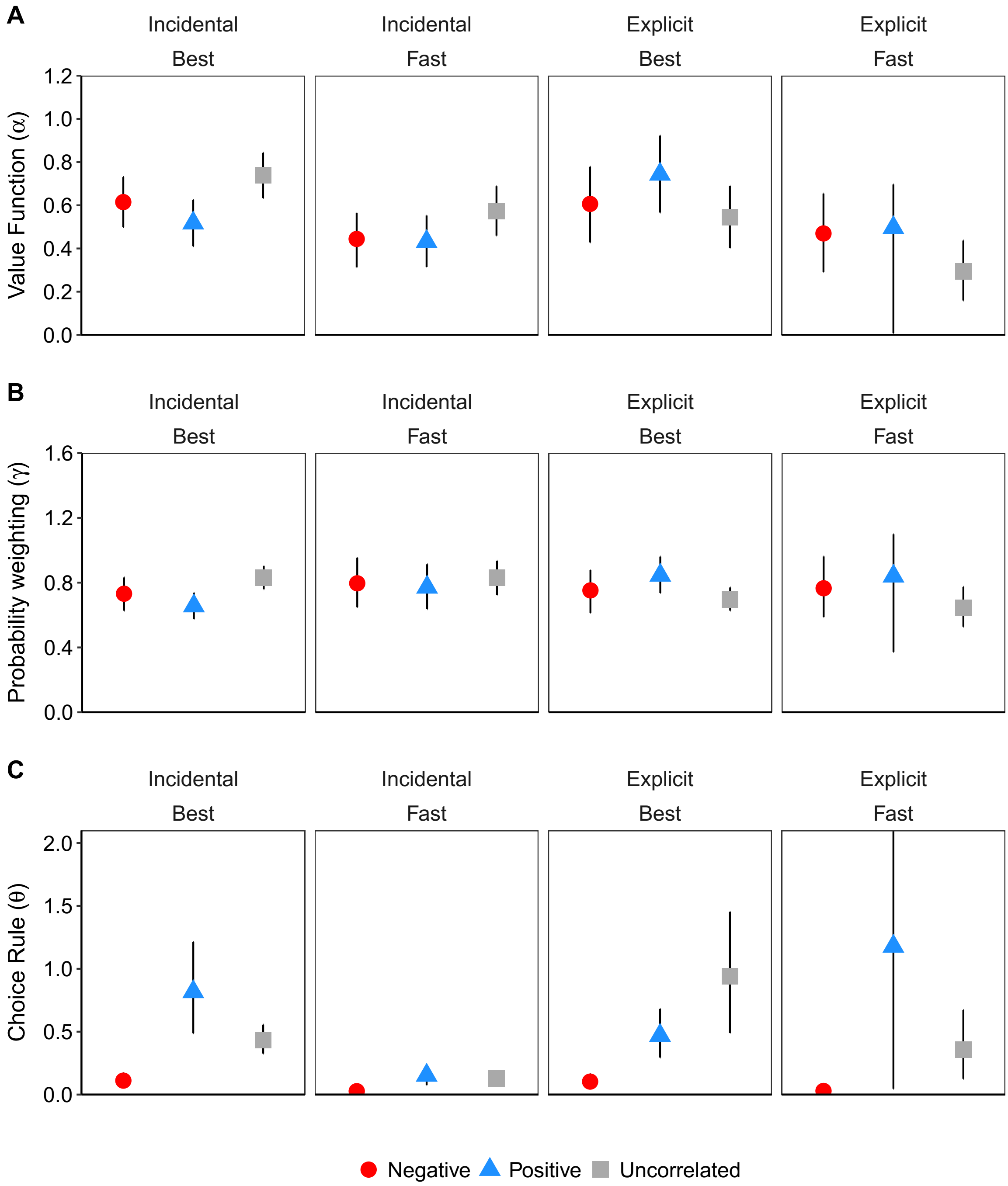


Figure C9. Posterior distributions for the group-level parameter estimates of Prospect Theory per condition.

References

Becker, G. M., Degroot, M. H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, 9(3):226–232.

Cavanagh, J. F., Wiecki, T. V., Kochar, A., and Frank, M. J. (2014). Eye Tracking and Pupillometry are Indicators of Dissociable Latent Decision Processes. *Journal of Experimental Psychology: General*, 143(4):1476–1488.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Klayman, J. (1988). On the How and Why (not) of Learning from Outcomes. In *Human Judgment: The SJT View*, volume 54, pages 115–162.

Leuker, C., Pachur, T., Hertwig, R., and Pleskac, T. J. (2018a). Exploiting risk–reward structures in decision making under uncertainty. *Cognition*, pages 186–200.

Leuker, C., Pachur, T., Hertwig, R., and Pleskac, T. J. (2018b). Too good to be true? Psychological responses to surprising options in risk–reward environments (Manuscript submitted for publication).

Luce, R. D. (1959). *Individual Choice Behavior*. John Wiley Sons, Inc., New York, NY.

McFadden, D., Machina, M. J., and Baron, J. (1999). Rationality for economists? In Fischhoff, B. and Manski, C. F., editors, *Elicitation of preferences*, pages 73–110. Springer, Dordrecht.

Olschewski, S., Rieskamp, J., and Scheibehenne, B. (2018). Taxing cognitive capacities reduces choice consistency rather than preference: A model-based test. *Journal of Experimental Psychology: General*, 147(4):462–484.

Peirce, J. W. (2007). Psychopy - Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1-2):8–13.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323.

Wattenmaker, W. D. (1991). Learning modes, feature correlations, and memory-based categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(5):908–923.

Whittlesea, B. W. (1987). Preservation of specific experiences in the representation of general knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(1):3–17.

Young, D. L., Goodie, A. S., Hall, D. B., and Wu, E. (2012). Decision making under time pressure, modeled in a prospect theory framework. *Organizational Behavior and Human Decision Processes*, 118(2):179–188.

D | Supplementary Material to Chapter 5: "When money talks: Judging risk and coercion in high-paying clinical trials"

1. Clinical trial vignette

Suppose that you are a member of an ethics committee, and you will have to decide whether or not to approve of the following study. Pay close attention. All the following questions will be based on this text.

The E.M.C.A. Medical Research Institute has developed a new vaccine to prevent infection with the Ebola virus. In rats and chimps the vaccine successfully prevents infection with the virus and causes no measurable side effects. The institute now seeks to enlist 100 female participants to investigate whether the vaccine causes side effects in women. This is important to know, as it will determine whether the vaccine can be given to female healthcare workers in regions affected by the disease.

Each of the 100 participants will be injected with the vaccine and then monitored in weekly intervals for two months. The total time required to participate if no side effects occur is about 40 hours. Participants will not be exposed to the virus; the study only tests for side effects of the vaccine. Since no side effects occurred in the animal studies, the institute's experts consider it unlikely that they will occur in humans. However, nobody knows for sure. This is why the experiment needs to be run. If unexpected side effects occur, they might range from very mild, such as a day of nausea, to very severe, such as persistent migraines. Side effects will be treated free of charge, if treating them is medically possible. An affected woman will not, however, receive treatment for any unrelated medical problems, and she will not receive any other compensation for suffering these side effects. The only compensation to any participant is the money paid to her when she agrees to take part in the study, before she is injected with the vaccine.

Study participation invitations will be posted throughout the city in which the institute is located. Invitations will be put up in both rich and poor neighborhoods. The institute will compensate each woman who participates with [50/1,000/10,000] for the risk she takes and the time commitment required to participate in the study (~ 40 hours).

How many participants would you expect to have *any* side effects from taking part in the study? 0-100 [slider]

How many participants would you expect to have *mild* side effects from taking part in the study? 0-100 [slider]

How many participants would you expect to have *severe* side effects from taking part in the study? 0-100 [slider]

NEW PAGE

[Reminder of the introductory text]

Suppose you are a member of the ethics committee that has to approve the institute's study with [payment 1]. How would you decide? [7-point Likert scale with extremes labelled "definitely reject" and "definitely approve"]

How much do you personally approve of the institute's proposal to enlist and compensate study participants from both rich and poor neighborhoods in this way? [7-point Likert scale with extremes labelled "strongly disapprove" and "approve without reservation"]

NEW PAGE

A.S. is a woman who lives in a poor part of the city. For the past 20 years she has worked in various minimum-wage jobs. She currently earns 1,500 per month, which is barely enough for her to get by. A.S. encounters one of the study participation invitations that the institute has posted on bulletin boards in her neighborhood. A.S. considers signing up for this study. She is on the fence about whether or not to do so. She is afraid of possible unexpected serious side effects of the vaccine. But then again, she would be paid [as much as she earns in her job in a day / almost as much as she earns in her usual job in an entire month / more than six times as much as in her usual job in an entire month].

NEW PAGE

[Reminder of the text on A.S.]

Suppose that 10 women similar to A.S. see the institute's study participation invitation. How many of the 10 would be better off if the institute had never posted the study participation invitation? [0-10]

How many of the 10, do you think, will eventually participate in the study in exchange for [payment 1]? [0-10]

If A.S. decides to participate in the study for [payment 1], how would you describe her decision? [7-point Likert scale with extremes labelled "she was coerced" and "her decision was entirely voluntary"]

If A.S. decides to participate in the study, how likely is it that she will later regret her decision? [7-point Likert scale with extremes labelled "extremely unlikely" and "extremely likely"]

If A.S. decides NOT to participate in the study, how likely is it that she will later regret her decision? [Answer choices: 7-point Likert scale with extremes labelled "extremely unlikely" and "extremely likely"]

Researchers at the institute discuss offering payment [2, 3] for participation instead. (Each participant saw the same questions again for the other two payment amounts. These data were used for within-respondent analyses.)

For each of the following ways of compensating study participants, please indicate how ethically appropriate you think it is. Recall that the study test for effects of a vaccine, and although nobody expects such side effects to occur, if this were known, there would be no need to run a study. Recall that there is no special compensation if side effects occur. [7-point Likert scale with extremes labelled "completely unethical" and "completely ethical"]

- Do not pay money for participation
- Pay 50 for participation
- Pay 1,000 for participation
- Pay 10,000 for participation
- Pay everyone the amount of money for participation that he would earn at his job in 40 hours

NEW PAGE

[Reminder of the text on A.S.]

For your preferred way of compensating participants, please briefly explain why you think it is the most ethical way to do it. [box in which participants could fill in an essay]

NEW PAGE

Have you ever participated in a medical research study?

Have you ever thought about participating in a medical research study as a means to earn money?

Would you participate in the experiment about the Ebola vaccine described in this study for a payment of 50?

Would you participate in the experiment about the Ebola vaccine described in this study for a payment of 1,000?

Would you participate in the experiment about the Ebola vaccine described in this study for a payment of 10,000? [“yes”, “no”, “I do not know”. In the first two of the above five questions, the choice “prefer not to answer” was also available.]

To what extent do you agree/disagree with the following statements (about clinical trial markets) in general? A “clinical trial market” refers to research institutions being able to offer monetary incentives to prospective participants to take part in a clinical trial, who can freely agree or disagree to take part. [7-point Likert scale with extreme labelled “completely agree” and “completely disagree”]

- Clinical trial markets are deplorable.
- Clinical trial markets should be banned.
- Clinical trial markets should be tightly monitored by the government.
- Clinical trial markets are morally permissible.

Note. These are almost identical materials as in Ambuehl et al. (2015). The differences were as follows. We (1) added the side effects question, (2) removed “midwestern” from the city description and expressed the payment in due to our primarily British sample, (3) in wave 2, added questions pertaining to how repugnant respondents considered clinical trial markets (see last question set), (4) asked respondents about their willingness to take risks and numeracy in addition to the other demographic variables (also see demographics table) and (5) appended an exploratory task in which respondents estimated their chances of loosing various hypothetical monetary amounts (see osf.io/5kewt/). We did not have predictions about how responses to this task are linked to our questions of interest here; the data will be analyzed and reported at a later point in time.

Ambuehl, S., Niederle, M., & Roth, A.E. (2015). More Money, More Problems? Can High Pay Be Coercive and Repugnant? *American Economic Review*, *105*(5), 357–360.

2. Clinical trial evaluations (replication)

Between-respondents

Variable	Doubters	Others	Trusters	×Doubter	×Doubter×wave
P (enroll)	-0.22 (-0.86; 0.41)	-0.08 (-0.71; 0.55)	0.16 (-0.24; 0.55)	-0.39 (-1.13; 0.37)	-0.22 (-2.01; 1.50)
Voluntariness	-0.78 (-1.26; -0.32	-0.43 (-0.89; 0.03)	0.18 (-0.08; 0.46)	-0.97 (-1.49; -0.44)	-0.72 (-1.96; 0.51)
P (better off)	0.75 (-0.14; 1.64)	-0.19 (-1.06; 0.68)	-0.50 (-1.00; 0.00)	1.26 (0.26; 2.26)	-1.24 (-3.49; 1.08)
P (regret accepting)	0.29 (-0.13; 0.71)	0.30 (-0.11; 0.72)	-0.25 (-0.49; 0.00)	0.53 (0.05; 1.02)	0.21 (-0.91; 1.33)
P (regret rejecting)	-0.09 (-0.50; 0.31)	0.32 (-0.05; 0.70)	0.31 (0.09; 0.54)	-0.41 (-0.85; 0.04)	0.25 (-0.78; 1.26)
Personal approval	-0.56 (-0.98; -0.13)	-0.40 (-0.83; 0.03)	0.23 (-0.01; 0.46)	-0.78 (-1.26; -0.31)	1.48 (0.38; 2.60)
IRB approval	-0.91 (-1.30; -0.51)	-0.69 (-1.11; -0.28)	0.32 (0.09; 0.55)	-1.22 (-1.69; -0.77)	1.28 (0.20; 2.35)
Obs. (wave 1)	76	64	176		
Obs. (wave 2)	302	274	536		
Obs. (total)	378	338	712		

Table D1. Effects of increasing payment from 1,000 to 10,000. Columns 2–4 display the beta coefficients per dependent variable separately for each respondent type. Column 5 displays the beta coefficients of the interaction effects between respondent type and compensation amount. "Truster" was used as a baseline; reported are the effects of Doubters compared to Trusters. Column 6 displays the beta coefficients of a three-way interaction between respondent type, compensation amount, and wave, using wave 1 as a baseline. All coefficients are based on modeling the effect of 10,000, using 1,000 as a baseline, and are reported with their 95% credible intervals. For each respondent, only the first payment amount they were presented with was included in the analysis (i.e., models are based on between-respondent analyses).

Note. The analyses show a credible difference between waves for the variables *Personal approval* and *IRB approval* (column 6). We investigated this result further by subsetting the data for each wave. This indicates the same direction of results but a noncredible effect in wave 2 for *Personal approval* (wave 1 = -1.85, CI = $[-2.90; -0.78]$; wave 2 = -0.37, CI = $[-0.90; 0.16]$; interaction effects between compensation amount and type using "Truster" as a baseline). Regarding *IRB approval*, the data were in the same direction and credible across both waves (wave 1 = -2.18, CI = $[-3.17; -1.21]$, wave 2 = -0.89, CI = $[-1.41; -0.36]$). As the interaction effects between compensation amount and type (using "Truster" as a baseline) were slightly weaker in wave 2, there was a credible interaction effect (see table).

Within-respondents

Variable	Doubters	Others	Trusters	×Doubter	×Doubter×wave
P (enroll)	1.40 (1.22; 1.57)	1.38 (1.19; 1.57)	1.72 (1.58; 1.86)	-0.32 (-0.54; -0.09)	-0.04 (-0.52; 0.44)
Voluntariness	-1.03 (-1.17; -0.90)	-0.18 (-0.29; -0.06)	-0.05 (-0.14; 0.05)	-0.98 (-1.14; -0.83)	-0.09 (-0.47; 0.29)
P (better off without)	0.20 (-0.06; 0.47)	-0.03 (-0.33; 0.27)	-0.55 (-0.77; -0.32)	0.75 (0.39; 1.12)	0.00 (-0.88; 0.88)
P (regret accepting)	-0.18 (-0.31; -0.04)	-0.33 (-0.45; -0.20)	-0.62 (-0.72; -0.52)	0.44 (0.28; 0.61)	-0.16 (-0.56; 0.23)
P (regret rejecting)	-0.18 (-0.31; -0.04)	-0.33 (-0.46; -0.21)	-0.62 (-0.72; -0.52)	0.45 (0.28; 0.61)	-0.17 (-0.57; 0.23)
Personal approval	-0.98 (-1.16; -0.82)	-0.04 (-0.17; 0.07)	0.54 (0.45; 0.63)	-1.52 (-1.70; -1.37)	0.09 (-0.31; 0.49)
IRB approval	-1.59 (-1.77; 1.40)	-0.30 (-0.46; -0.15)	0.81 (0.69; 0.93)	-2.41 (-2.61; -2.19)	-0.19 (-0.69; 0.31)

Table D2. Replication of Table D1, but including all three payment amounts per respondent (within-respondent analyses). We accounted for random variation between respondents by including Response ID as a grouping factor.

3. Side effects

Between-respondents

Side effect type	Doubters	Others	Trusters	×Doubter	×Doubter×wave
Any	6.19 (1.15; 11.33)	-2.15 (-8.32; 4.01)	-0.91 (-4.48; 2.71)	7.08 (0.35; 13.77)	6.82 (-8.59; 22.54)
Mild	7.47 (3.30; 11.46)	-0.73 (-5.82; 4.48)	-0.09 (-2.92; 2.71)	7.52 (2.75; 12.96)	-0.01 (-12.62; 12.55)
Severe	3.00 (-0.15; 6.17)	-2.45 (-6.11; 1.23)	1.57 (-0.70; 3.84)	1.30 (-2.63; 5.30)	3.93 (-5.18; 13.14)

Table D3. Effects of increasing payment from 1,000 to 10,000. Columns 2–4 display the beta coefficients per dependent variable separately for each respondent type. Column 5 displays the beta coefficients of the interaction effects between respondent type and compensation amount. "Truster" was used as a baseline, reported are the effects of Doubters compared to Trusters. Column 6 displays the beta coefficients of a three-way interaction between respondent type, compensation amount and wave, using wave 1 as a baseline. All coefficients are based on modeling the effect of 10,000, using 1,000 as a baseline, and are reported with their 95% credible intervals. For each respondent, only the first payment amount they were presented with was included in the analysis (i.e., models are based on between-respondent analyses).

Within-respondents

Side effect type	Doubters	Others	Trusters	×Doubter	×Doubter×wave
Any	3.19 (1.96; 4.42)	2.00 (0.43; 3.61)	1.61 (0.75; 2.46)	1.58 (0.00; 3.15)	-2.35 (-6.12; 1.43)
Mild	2.03 (0.85; 3.20)	1.85 (0.47; 3.22)	0.77 (0.14; 1.41)	1.26 (-0.07; 2.57)	-2.37 (-5.62; 0.91)
Severe	2.49 (1.59; 3.38)	0.84 (-0.21; 1.90)	1.17 (0.57; 1.78)	1.32 (0.24; 2.43)	-2.72 (-5.40; -0.03)

Table D4. Replication of Table D3, but including all three payment amounts per respondent (within-respondent analyses). We accounted for random variation between respondents by including Response ID as a grouping factor. The analysis reveals that Trusters slightly adjusted the estimated number of side effects upwards as payoffs became very large; the effect was more pronounced for Doubters (comparison of b regressors and interaction effect in column 5).

Note. We also tested for order effects. It is plausible that within-respondents effects would be larger when going from low to high payment amounts. Indeed, the within-respondent effects were mostly driven by respondents who were presented with 50 first, and the larger amounts later ($b_{Trusters}$ = 2.63, CI = $[1.04, 4.24]$, $b_{Doubters}$ = 3.74, CI = $[1.89, 5.60]$, estimated number of any side effects when 50 was presented first). Because of this, we relied on between-respondent differences in our primary analyses.

4. Combined models

Clinical trial evaluations (regressions)

Variables	Side eff. any	S&V: Side effects	S&V: Voluntariness	Side eff. mild	Side eff. severe
P (enroll)	0.0017 (-0.0053; 0.0087)	0.0031 (-0.0040; 0.0102)	0.1190 (0.0290; 0.2048)	0.0000 (-0.0088; 0.0088)	0.0040 (-0.0084; 0.0163)
Voluntariness	–	–	–	–	–
P (better off without)	0.0339 (0.0248; 0.0429)	0.0299 (0.0209; 0.0390)	-0.3709 (-0.4819; -0.2594)	0.0448 (0.0334; 0.0563)	0.0548 (0.0388; 0.0709)
P (regret accepting)	0.0216 (0.0173; 0.0259)	0.0188 (0.0146; 0.0229)	-0.2612 (-0.3118; -0.2107)	0.0264 (0.0210; 0.0317)	0.0284 (0.0208; 0.0361)
P (regret rejecting)	-0.0068 (-0.0110; -0.0026)	-0.0063 (-0.0105; -0.0021)	0.0495 (-0.0021; 0.1017)	-0.0081 (-0.0132; -0.0029)	-0.0142 (-0.0214; -0.0069)
Personal approval	-0.0087 (-0.0132; -0.0042)	-0.0054 (-0.0097; -0.0011)	0.3051 (0.2522; 0.3581)	-0.0107 (-0.0163; -0.0051)	-0.0125 (-0.0201; -0.0048)
IRB approval	-0.0129 (-0.0173; -0.0086)	-0.0098 (-0.0139; -0.0057)	0.2800 (0.2291; 0.3311)	-0.0138 (-0.0193; -0.0084)	-0.0172 (-0.0147; -0.0096)
$\delta_{IRB\ approval}$*	-0.0190 (-0.0311; -0.0069)	-0.0180 (-0.0303; -0.0059)	0.0959 (-0.0561; 0.2472)	-0.0241 (-0.0391; -0.0090)	-0.0229 (-0.0444; -0.0012)
$\delta_{personal\ approval}$*	-0.0070 (-0.0161; 0.0023)	-0.0072 (-0.0166; 0.0021)	-0.0074 (-0.1233; 0.1078)	-0.0077 (-0.0190; 0.0037)	-0.0119 (-0.0279; 0.0043)
$\delta_{ethicality}$*	-0.0068 (-0.0084; 0.0048)	-0.0042 (-0.0158; 0.0074)	0.2634 (0.1187; 0.4070)	-0.0159 (-0.0307; -0.0009)	-0.0013 (-0.0216; 0.0190)

Table D5. Results for models linking clinical trial evaluations to respondents' estimated number of side effects. Columns 2–4 use the variable "any" side effects as the predictor. Column 2 shows the main effect; columns 3 and 4 show β weights from a model that included both side effects and voluntariness as predictors ("S&V", modeled as two main effects). Columns 5 and 6 show main effects using "mild" and "severe" side effects as predictors.

*same models as above, but including an interaction with payoff (using $1,000$ as a baseline) — as the δs also refer to payoff–dependent differences in the evaluations.

Note. We also tested a number of other models, for instance controlling for Doubter/Truster × compensation amount, the interaction between side effects estimates and payment amount, or the three-way interaction Doubter/Truster × compensation amount × side effects estimates. The main effects of side effect estimates on the clinical trial evaluations was present in all models. Generally, and as can be seen from Table D5, a higher number of estimated side effects results in a less positive evaluation of a medical trial, and a higher degree of "voluntariness" for prospective participants results in a more positive evaluation of a medical trial.

Clinical trial evaluations (mediations)

We used a mediation approach suggested by MacKinnon et al. (2000) to show how payment information differentially affected IRB approval of Doubters and Trusters, given their differential estimates of side effects and coerciveness.

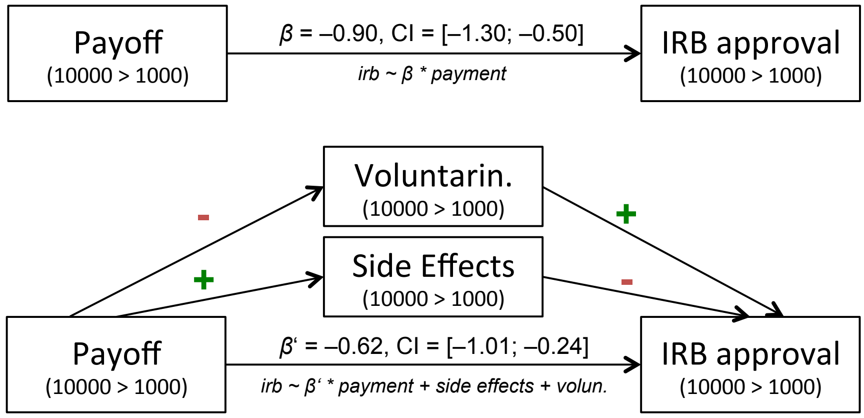


Figure D1. Doubters. Voluntariness and estimated side effects **lower** the role payoff plays in determining IRB approval (mediation effect: $\beta - \beta'$).

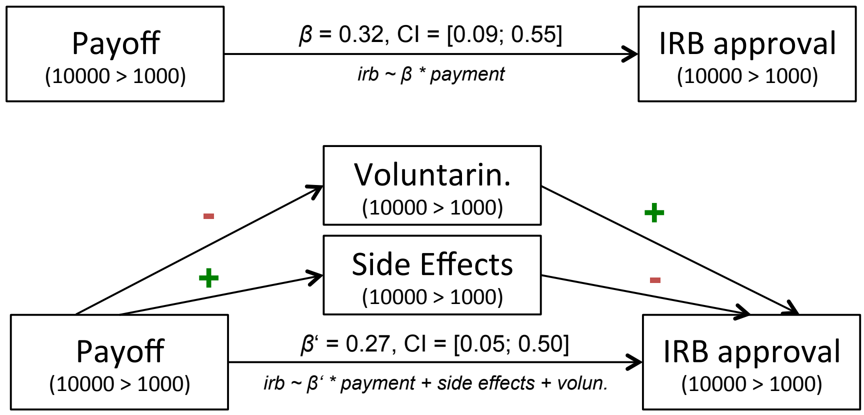


Figure D2. Trusters. Voluntariness and estimated side effects **do not (strongly) affect** the role payoff plays in determining IRB approval substantially (mediation effect: $\beta - \beta'$).

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding, and suppression effect. *Prevention Science* (1), 173–181.

Individual payoff sensitivity

Our findings suggest that Doubters (1) were more likely to utilize payment amount as a cue when estimating side effects; (2) also considered high payments to increase coercion (decrease voluntariness); and (3) evaluated high payments as less ethical (IRB approval). Is there a link between sensitivity to payoff information for IRB approval and side effects; and/or for IRB approval and voluntariness? In an exploratory analysis, we turned to within-respondent analyses for these dependent variables (i.e., using respondents' answers pertaining to each payment amount). For each respondent, we computed a "payoff sensitivity score" that measured how responses changed as clinical trials offered extremely high compensations. We did this for respondents' clinical trial evaluations (IRB approval: $\delta_{IRB} = IRB_{10,000} - IRB_{1,000}$), for their inferred number of any side effects [SE] ($\delta_{SE} = SE_{10,000} - SE_{1,000}$), and for voluntariness ($\delta_{volun.} = volun._{10,000} - volun._{1,000}$).

As Figure D3 shows, δ_{IRB} and δ_{SE} were inversely related for Doubters ($b = -0.018$, CI = $[-0.032, -0.004]$) but not for Trusters ($b = -0.006$, $CI = [-0.017, 0.004]$). In addition, δ_{IRB} and $\delta_{voluntariness}$ were positively related for both Doubters ($b = 0.246$, CI = $[0.110, 0.382]$), and Trusters ($b = 0.151$, CI = $[0.118, 0.243]$; $\delta_{voluntariness}$ not plotted). A comparison of β coefficients suggests that the link was stronger for Doubters, again suggesting a stronger reliance on payoff information when evaluating clinical trials.

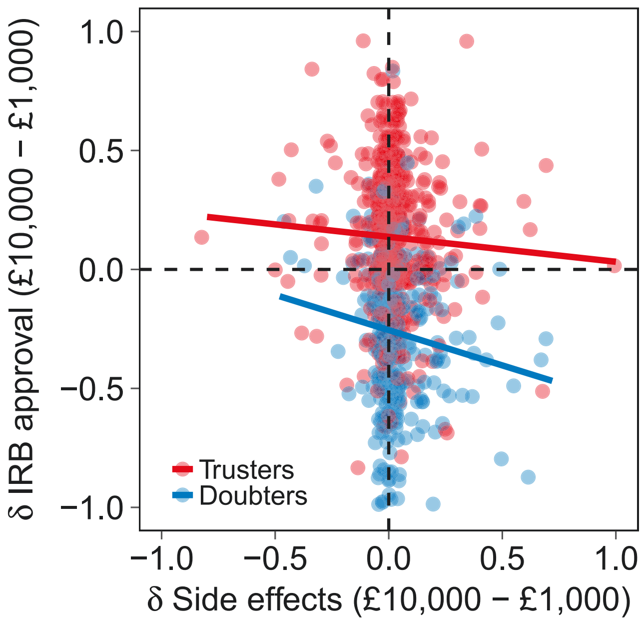


Figure D3. Relation between payoff sensitivity in the estimated number of side effects (normalized δ_{SE}) and payoff sensitivity in IRB approval ratings (normalized δ_{IRB}). A payoff-dependent increase in estimated side effects was linked to lower IRB approval rates for Doubters, but not Trusters (within-respondents). Each dot represents one participant. Note: Although the relationship is statistically reliable, as the density of the datapoints suggests, most respondents did not infer a different amount of side-effects—which is sensible given they saw the same vignette (just with a different payoff) repeatedly.

5. Repugnance

	Side eff.	S&V: Side eff.	S&V: Volunt.	Side eff. (50)	Side eff. (1000)
Repugnance	0.0082 (0.0058; 0.0106)	0.0065 (0.0042; 0.0088)	-0.1489 (-0.1803; -0.1172)	0.0089 (0.0062; 0.0116)	0.0090 (0.0063; 0.0116)

Table D6. Columns 2–5 used participants estimates of "any" side effects for 10,000 as predictors. Column 2 shows the main effect; columns 3 and 4 show β weights from a model that included both side effects and voluntariness as predictors ("S&V", modeled as two main effects). Columns 5 and 6 show main effects using side effects estimates for 50 and 1,000 as predictors.

	Side eff. Doubters	Side eff. Others	Side eff. Trusters
Repugnance	0.0082 (0.0039; 0.0125)	0.0142 (0.0092; 0.0194)	0.0048 (0.0013; 0.0082)

Table D7. Side effects. Main effects for each type. Each respondent is only entered in the regression once, with their estimate for "any' side effects given the 10,000 trial.

	Volunt. Doubters	Volunt. Others	Volunt. Trusters
Repugnance	-0.1240 (-0.1805; -0.0675)	-0.2386 (-0.3073; -0.1705)	-0.1531 (-0.1998; -0.1060)

Table D8. Voluntariness. Main effects for each type. Each respondent is only entered in the regression once, with their estimate for "any' side effects given the 10,000 trial.

	Repugnance
Side eff.	0.0063 (0.0039; 0.0087)
Voluntariness	-0.1606 (-0.1934; -0.1281)
Risk (health)	-0.0328 (-0.0598; -0.0058)
Income	-0.0701 (-0.1349; -0.0018)
Gender (male)	-0.1072 (-0.2172; 0.0033)
Thought about participating	-0.2524 (-0.5000; -0.0069)
Doubter	-0.1071 (-0.2340; 0.0214)

Table D9. Repugnance (column 2) predicted from side effects, voluntariness, and demographic variables (also see S5 for main effects). As before, the combined model reveals that a higher number of estimated side effects also increases how repugnant clinical trials are considered; lower voluntariness (i.e., higher coercion) has the opposite effect. Beyond these predictors, only whether or not respondents had thought about participating themselves and willingness to take health risks predicted additional, unique variance. Being a Doubter was not a reliable predictor (CI includes 0).

6. Demographics

Variables	$\delta_{\,IRB\ approval}$	$\delta_{\,personal\ approval}$	$\delta_{\,ethicality}$	Repugnance
Income	-0.29 (-0.42; -0.15)	-0.07 (-0.18; 0.03)	-0.20 (-0.34; -0.07)	-0.09 (-0.16; -0.01)
Education	-0.09 (-0.19; 0.00)	-0.07 (-0.09; 0.00)	-0.22 (-0.32; -0.13)	-0.06 (-0.11; -0.01)
Gender (male)	-0.16 (-0.36; 0.05)	-0.01 (-0.17; 0.05)	-0.01 (-0.20; 0.22)	-0.20 (-0.32; -0.08)
Age	-0.020 (-0.029; -0.012)	-0.004 (-0.010; 0.003)	-0.012 (-0.021; -0.004)	-0.006 (-0.010; -0.001)
Numeracy	-0.25 (-0.53; 0.04)	-0.12 (-0.33; 0.10)	-0.50 (-0.78; -0.22)	-0.15 (-0.29; 0.02)
Risk-taking (health)	0.00 (-0.05; 0.05)	-0.02 (-0.02; 0.02)	-0.05 (-0.08; -0.02)	-0.05 (-0.08; -0.02)
Risk-taking (general)*	0.05 (-0.02; 0.11)	0.01 (-0.04; 0.06)	0.08 (0.01; 0.14)	-0.01 (-0.04; 0.02)
Thought about participating	-0.06 (-0.28; 0.15)	-0.11 (-0.27; 0.05)	-0.10 (-0.32; 0.11)	-0.26 (-0.38; -0.15)
$\delta_{\,IRB\ approval}$				
$\delta_{\,personal\ approval}$	0.75 (0.70; 0.81)			
$\delta_{\,ethicality}$	0.57 (0.53; 0.61)	0.36 (0.33; 0.40)		
Repugnance (wave 2)	0.119 (-0.005; 0.244)	0.091 (-0.004; 0.185)	0.080 (-0.043; 0.203)	

Table D10. Demographic variables and relationship to measures of interest. δs are the differences in respondents' evaluations of the clinical trial offering 10,000 vs. 1,000 ($\delta = Rating_{10,000} - Rating_{1,000}$). All coefficients are reported with their 95% credible intervals. As in the original survey, we find that personal approval and IRB approval are highly correlated ($\beta = .76$). Moreover, IRB approval and ethicality are correlated, but to a lesser extent ($\beta = .36$). The distinction between Doubters and Trusters is based on $\delta_{\,ethicality}$.

Declaration of Independent Work

Portions of this dissertation have been prepared in collaboration with co–authors for publication. I am the primary contributor in all regards. For the most up to date versions, see the following:

Chapter 2 has been published in Cognition, and as a preprint on the Open Science Framework, at osf.io/tmcnd/. Chapter 3 is available as a preprint on the Open Science Framework, at osf.io/xautn/. Chapter 4 is currently in preparation, a preprint will be added to the project's preregistration on the Open Science Framework, at osf.io/pxt7a/. Chapter 5 is available as a preprint on the Open Science Framework, at osf.io/9wmta/. All preprints have an Attribution 4.0 International (CC BY 4.0) license and can thus be reprinted.

I hereby declare that:

- I completed the doctoral thesis independently. Except where otherwise stated, I confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.
- I have not applied for a doctoral degree elsewhere and do not have a corresponding doctoral degree.
- I have not submitted the doctoral thesis, or parts of it, to another academic institution and the thesis has not been accepted or rejected.
- I have acknowledged the Doctoral Degree Regulations which underlie the procedure of the Faculty of Life Sciences of Humboldt-Universität zu Berlin, as amended on 5th March 2015.
- No collaboration with commercial doctoral degree supervisors took place.
- The principles of Humboldt-Universität zu Berlin for ensuring good academic practice have been complied with.

Christina Leuker
Berlin, July 2018